

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 00:03:32 ; Search time 9715 Seconds  
(without alignments)  
11591.346 Million cell updates/sec

Title: US-09-436-184-3

Perfect score: 2324

Sequence: 1 cggacgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_rcg.\*  
3: gb\_in.\*  
4: gb\_cm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	6	AX146795 Sequence
2	2324	100.0	2324	9	S83325 Sequence
3	2320.8	99.9	2452	6	CQ871564 Sequence
4	2296.8	98.8	2449	9	HSU03109 Sequence
5	1480.4	63.7	2739	4	BOVASHY
6	1450.2	62.4	4419	10	AF289486
7	1450.2	62.4	6629	10	AF289487
8	1215.8	52.3	1970	6	CQ729666
9	947	40.7	2680	9	AF289489
10	829.6	35.7	1080	9	BC015518
11	819	35.2	2442	6	BD234762
12	819	35.2	2442	6	AR274038
13	819	35.2	2442	6	AR277619
14	819	35.2	2442	6	AR441744
15	819	35.2	2442	6	AR532049
16	819	35.2	2442	6	AX369081
17	779.4	33.5	2280	9	AF306765
18	679	29.2	2761	9	BC025236
19	455.8	19.6	2789	10	AF289488

20	412.2	17.7	2617	10	AF302653	Mus muscu
21	400.8	17.2	2770	10	BC015281	Mus muscu
22	364.4	15.7	366	6	CQ680735	Sequence
23	297	12.8	297	6	CQ731371	Sequence
24	279.2	12.0	1608	9	BC066929	Homo sapi
25	275.8	11.9	2637	10	AF302654	Mus muscu
26	267.8	11.5	524	10	AF289491	Homo sapi
27	220.2	9.5	855	9	AF224468	Mus muscu
28	218	9.4	2777	3	AF289493	Drosophil
29	215.8	9.3	566	10	AF289492	Mus muscu
30	215.8	9.3	988	10	AF223414	Mus muscu
31	215.4	9.3	597	5	BX931169	Sequence
32	215.4	9.3	652	5	CR406273	Gallus ga
33	213.6	9.2	1662	6	CQ584991	Sequence
34	208.6	9.0	3170	3	AF289494	Drosophil
35	199.2	8.6	3237	3	BT003633	Drosophil
36	195.6	8.4	1014	5	BX929679	Gallus ga
37	194.2	8.4	1777	4	CFU38414	Canis fami
38	190	8.2	132868	2	AC021323	Homo sapi
39	190	8.2	158404	9	AC090094	Homo sapi
40	177.6	7.6	2432	10	AF302655	Mus muscu
41	173.8	7.5	919	10	AF221854	Mus muscu
42	173.8	7.5	1013	10	AF223413	Mus muscu
43	173.8	7.5	1733	10	AF289490	Mus muscu
44	165.2	7.1	941	9	AF224469	Homo sapi
45	165.2	7.1	1254	9	AF184241	Homo sapi

## ALIGNMENTS

RESULT 1  
AX146795  
LOCUS AX146795 2324 bp DNA linear PAT 31-MAY-2001  
DEFINITION Sequence 3 from Patent WO0135102.  
ACCESSION AX146795  
VERSION AX146795.1 GI:14285147  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
Wands, J.R., de la Monte, S.M., Ince, N. and Carlson, R.I.  
Diagnosis and treatment of malignant neoplasms  
Patent: WO 0135102-A 3 17-MAY-2001;  
RHODE ISLAND HOSPITAL (US)

FEATURES  
source  
1..2324  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN	Query Match	Best Local	Similarity	Score	DB 6;	Length	2324;
	Matches 2324;	Conservative	0;	Mismatches	0;	Indels	0;
QY	1	CGGACCGTGCATGGCCAGCGTAAAGATGCCAAGAGCAGCGCGCAACAGCAGCAGCGG	60				
DB	1	CGGACCGTGCATGGCCAGCGTAAAGATGCCAAGAGCAGCGCGCAACAGCAGCAGCGG	60				
QY	61	GCTCCGCGAGCGGTAGCAGTGGCGGACAGCAGCCCCCGGGCCCGGAGAGACAA	120				
DB	61	GCTCCGCGAGCGGTAGCAGTGGCGGACAGCAGCCCCCGGGCCCGGAGAGACAA	120				
QY	121	AGCATGGAGCACACAGATGGAGGAAGCGGACTCTCGGAACTTCATTCTTCAGT	180				
DB	121	AGCATGGAGCACACAGATGGAGGAAGCGGACTCTCGGAACTTCATTCTTCAGT	180				
QY	181	GGTTTATGGTATGTCATTCGTGGCGCTGTGGACATCTGTAGCTGCTGTTTGTGATC	240				
DB	181	GGTTTATGGTATGTCATTCGTGGCGCTGTGGACATCTGTAGCTGCTGTTTGTGATC	240				

QY	241	TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGGTATGGAG	300
Db	241	TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGGTATGGAG	300
QY	301	ATTTTGATGTGGATGATGCCAAAGTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360
Db	301	ATTTTGATGTGGATGATGCCAAAGTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360
QY	361	CAGCAGTCCCAGAGAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGTTCTCTGTGG	420
Db	361	CAGCAGTCCCAGAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGTTCTCTGTGG	420
QY	421	AGGCAGAACCCAGAAATATCAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATG	480
Db	421	AGGCAGAACCCAGAAATATCAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATG	480
QY	481	AAATGGTACACGCGAGAAATCTTGGGAGAGAGATGTCGCAACAAAGAGATGGACCCACAG	540
Db	481	AAATGGTACACGCGAGAAATCTTGGGAGAGAGATGTCGCAACAAAGAGATGGACCCACAG	540
QY	541	GAGAACCAACAGAGATGATGAGTTTCTATGAGGAGAAAGTTCGCAACAAAGAGATGGACCCACAG	600
Db	541	GAGAACCAACAGAGATGATGAGTTTCTATGAGGAGAAAGTTCGCAACAAAGAGATGGACCCACAG	600
QY	601	AGACCTGGAACCTGAACTATCTCATGAAGAAACCGAGCATAGTTACCACTGTGGAAGAGA	660
Db	601	AGACCTGGAACCTGAACTATCTCATGAAGAAACCGAGCATAGTTACCACTGTGGAAGAGA	660
QY	661	CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG	720
Db	661	CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG	720
QY	721	ATTCAGTGAACAGTATGAGAGTGAAGATGACCACTGACCATGATCATGTAAAT	780
Db	721	ATTCAGTGAACAGTATGAGAGTGAAGATGACCACTGACCATGATCATGTAAAT	780
QY	781	ACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA	840
Db	781	ACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA	840
QY	841	CAGAAATGCTCTCCCTCAGGATATCTGTTAGAAATGATTCAGAGTAAATTTAGAG	900
Db	841	CAGAAATGCTCTCCCTCAGGATATCTGTTAGAAATGATTCAGAGTAAATTTAGAG	900
QY	901	AGTAAAGCATTTTCTGTTGGAAGAACAGCAGGAGTACCCAGAAACAAATAGAAAAA	960
Db	901	AGTAAAGCATTTTCTGTTGGAAGAACAGCAGGAGTACCCAGAAACAAATAGAAAAA	960
QY	961	CAGATGATCCAGAACAAAAAGTAAAGTAAAGAAAGAGCCTAAACTTTTAAATAAAT	1020
Db	961	CAGATGATCCAGAACAAAAAGTAAAGTAAAGAAAGAGCCTAAACTTTTAAATAAAT	1020
QY	1021	TTGATAAGCATTTTAAAGTGAATTTGATGCTGCGAAAAAACTCCGTTAAAGGGAAAAA	1080
Db	1021	TTGATAAGCATTTTAAAGTGAATTTGATGCTGCGAAAAAACTCCGTTAAAGGGAAAAA	1080
QY	1081	TTGAGAGACGATGATTTAAAGAACTAGTACGCAATACCTCAGAGTCCACGAG	1140
Db	1081	TTGAGAGACGATGATTTAAAGAACTAGTACGCAATACCTCAGAGTCCACGAG	1140
QY	1141	CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGTAATGAGG	1200
Db	1141	CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGTAATGAGG	1200
QY	1201	TGCTAGTGGAGCATCGAGACCTACCAAGAGGTGGCCAGCTCCTGATGTCCTGCGAG	1260
Db	1201	TGCTAGTGGAGCATCGAGACCTACCAAGAGGTGGCCAGCTCCTGATGTCCTGCGAG	1260
QY	1261	ACTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAAATTTCTAGTGCATATGA	1320
Db	1261	ACTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAAATTTCTAGTGCATATGA	1320

RESULT 2

S83325

LOCUS

DEFINITION

S83325 2324 bp mRNA linear PRI 28-MAR-1997  
aspartyl (asparaginyl) beta-hydroxylase [human, hepatoblastoma cell]

QY	1321	GAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCAACTATTTCCTCAATGATCTTCTCTTAA	1380
Db	1321	GAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCAACTATTTCCTCAATGATCTTCTCTTAA	1380
QY	1381	AAATGACCTTTGGCTGGGATACCTCTGATAGGAGATATGACAAATGCAAGAAAAGTTT	1440
Db	1381	AAATGACCTTTGGCTGGGATACCTCTGATAGGAGATATGACAAATGCAAGAAAAGTTT	1440
QY	1441	ATCAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA	1500
Db	1441	ATCAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA	1500
QY	1501	TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGATAGAAAT	1560
Db	1501	TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGATAGAAAT	1560
QY	1561	CCGAGATCTGCGACCTGATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGCAGA	1620
Db	1561	CCGAGATCTGCGACCTGATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGCAGA	1620
QY	1621	GGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGAGACACTTTG	1680
Db	1621	GGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGAGACACTTTG	1680
QY	1681	CATCTGTCTGGCAACGCTCACTCAATGTGAATGGACTGGAAGCAGCAGCTTGGTGA	1740
Db	1681	CATCTGTCTGGCAACGCTCACTCAATGTGAATGGACTGGAAGCAGCAGCTTGGTGA	1740
QY	1741	CCCCAAAGAAACCGGCTCACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
Db	1741	CCCCAAAGAAACCGGCTCACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
QY	1801	TCCGAGATGAAGGCTTTGCAAGTATGATGAATAAAGCCAAAGAGTCTTCTCCTCGCTGAGATG	1860
Db	1801	TCCGAGATGAAGGCTTTGCAAGTATGATGAATAAAGCCAAAGAGTCTTCTCCTCGCTGAGATG	1860
QY	1861	AAAACTGAGGAAAAAGGGGACTGGAGCCAGTTCAGCTGTGGCAGCAAGGAGAGAGAA	1920
Db	1861	AAAACTGAGGAAAAAGGGGACTGGAGCCAGTTCAGCTGTGGCAGCAAGGAGAGAGAA	1920
QY	1921	ATGAAAATGCTGCAAAAGGAGCTCTAAAACCTGTACCTTACTAGAAAAGTTCCCCGAGA	1980
Db	1921	ATGAAAATGCTGCAAAAGGAGCTCTAAAACCTGTACCTTACTAGAAAAGTTCCCCGAGA	1980
QY	1981	CAACAGGATGAGAAAGAGGACAGATCAAAATTTCCATCATGACACCCCGGAGCTCACGTGT	2040
Db	1981	CAACAGGATGAGAAAGAGGACAGATCAAAATTTCCATCATGACACCCCGGAGCTCACGTGT	2040
QY	2041	GGCCGACACAGGCCCCACAACTGCGAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA	2100
Db	2041	GGCCGACACAGGCCCCACAACTGCGAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA	2100
QY	2101	AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACAGGACCTGGGAGGAAGGCAAGTGC	2160
Db	2101	AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACAGGACCTGGGAGGAAGGCAAGTGC	2160
QY	2161	TCATCTTTGATGACTCCTTTGAGCAGGATGATGGAGGATGCTCTCATCTTCCGGCTGA	2220
Db	2161	TCATCTTTGATGACTCCTTTGAGCAGGATGATGGAGGATGCTCTCATCTTCCGGCTGA	2220
QY	2221	TATTCATCGTGGATGTGGCATCCGAACTGACACACACAGCAGACGACGCTTCCAG	2280
Db	2221	TATTCATCGTGGATGTGGCATCCGAACTGACACACACAGCAGACGACGCTTCCAG	2280
QY	2281	CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA	2324
Db	2281	CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA	2324



1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGGAGATATGCAATGCAAGAAAGTTT 1440  
1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGGAGATATGCAATGCAAGAAAGTTT 1440  
1441 ATGAAGAGGCTGCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATATGCTTCA 1500  
1441 ATGAAGAGGCTGCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATATGCTTCA 1500  
1501 TCCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCCATATTTAAAGGAAGATAGAT 1560  
1501 TCCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCCATATTTAAAGGAAGATAGAT 1560  
1561 CCGGAGATCTCTGGCACTGATGATGGGAGATTTATTTCCACTGGGGATGCCATGCGA 1620  
1561 CCGGAGATCTCTGGCACTGATGATGGGAGATTTATTTCCACTGGGGATGCCATGCGA 1620  
1621 GGGTTGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGGACACTTTG 1680  
1621 GGGTTGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGGACACTTTG 1680  
1681 CATCTGTCTGCAACGCTCACTTACATGTGATGACTGAAGACAGCCTTTGGTGA 1740  
1681 CATCTGTCTGCAACGCTCACTTACATGTGATGACTGAAGACAGCCTTTGGTGA 1740  
1741 CCCCCAAAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTAA 1800  
1741 CCCCCAAAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTAA 1800  
1801 TCCGAGATGAAGGCTTCCAGTGTGATGATGAAGCAAGGCTCTCTCTGCTGAGGATG 1860  
1801 TCCGAGATGAAGGCTTCCAGTGTGATGATGAAGCAAGGCTCTCTCTGCTGAGGATG 1860  
1861 AAAACCTGAGGAAAGAGGACTGGAGCCAGTTTCAAGCTGTGACAGCAAGAGAAAGAA 1920  
1861 AAAACCTGAGGAAAGAGGACTGGAGCCAGTTTCAAGCTGTGACAGCAAGAGAAAGAA 1920  
1921 ATGAAATGCTGCAAGAGGACTCTTAAACCTGTACTTACTAGAAAGTTTCCCGAGA 1980  
1921 ATGAAATGCTGCAAGAGGACTCTTAAACCTGTACTTACTAGAAAGTTTCCCGAGA 1980  
1981 CAACAGGATGAGAGGACAGATCAATATTCATCATGCAACCCGGGACTCACGTGT 2040  
1981 CAACAGGATGAGAGGACAGATCAATATTCATCATGCAACCCGGGACTCACGTGT 2040  
2041 GGCGGCACACAGGCGCCCAAACTGCGAGCTCCGAATGCACTGGCTTGGTATTTCCCA 2100  
2041 GGCGGCACACAGGCGCCCAAACTGCGAGCTCCGAATGCACTGGCTTGGTATTTCCCA 2100  
2101 AGGAAGGCTGCAAGATTCTGATGTGCAACGAGACAGGACCTGGGAGGAAGGCAAGTGC 2160  
2101 AGGAAGGCTGCAAGATTCTGATGTGCAACGAGACAGGACCTGGGAGGAAGGCAAGTGC 2160  
2161 TCATCTTTGATGACTCTTTGAGCAAGGATGAGGAGGATGCTCATCTTTCCGGGTGA 2220  
2161 TCATCTTTGATGACTCTTTGAGCAAGGATGAGGAGGATGCTCATCTTTCCGGGTGA 2220  
2221 TATTTCATGCTGGATGTGGCATCCGGAACCTGACACACAGAGGACGAGCCTTCCAG 2280  
2221 TATTTCATGCTGGATGTGGCATCCGGAACCTGACACACAGAGGACGAGCCTTCCAG 2280  
2281 CAATTTAGCATGAATTTCAATGCAAGCTTGGGAACTCTGGAGAGA 2324  
2281 CAATTTAGCATGAATTTCAATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 3  
LOCUS CO871564  
DEFINITION Sequence 55 from Patent WO2004078783.  
ACCESSION CO871564  
VERSION CO871564.1 GI:52745624  
KEYWORDS linear PAT 27-SEP-2004

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Murphy, F.J., Sheehan, D.E., Keating, K.E., Hayes, I.E. and Seera, L.E.  
TITLE Enzymes involved in apoptosis  
JOURNAL Patent: WO 2004078783-A 55 16-SEP-2004;  
Biox Therapeutics Ltd (IE)  
FEATURES  
source Location/Qualifiers  
1..2452  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 99.9%; Score 2320.8; DB 6; Length 2452;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATATGGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 67 CGGACCGTGCATATGGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 126  
Qy 61 GCTCCGGCAGCGTAGCAGAGTGGCGGCGAGCAGCGCCCGGGGCGCCGGAGAGAGCAA 120  
Db 127 GCTCCGGCAGCGTAGCAGAGTGGCGGCGAGCAGCGCCCGGGGCGCCGGAGAGAGCAA 186  
Qy 121 AGCATGGAGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGAACTTCACTTTCACT 180  
Db 187 AGCATGGAGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGAACTTCACTTTCACT 246  
Qy 181 GGTATTATGATGATTCATCTCTGGCGCTCTGGGACATCTGTAGCTGTCTGTTGTTGATC 240  
Db 247 GGTATTATGATGATTCATCTCTGGCGCTCTGGGACATCTGTAGCTGTCTGTTGTTGATC 306  
Qy 241 TTGTTGATATGAGGAAGTTCTTAGGAAAATCTAGGAATCTATGATCTGTGATGATGATGAG 300  
Db 307 TTGTTGATATGAGGAAGTTCTTAGGAAAATCTAGGAATCTATGATCTGTGATGATGATGAG 366  
Qy 301 ATTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 367 ATTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 426  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTCTG 420  
Db 427 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTCTG 486  
Qy 421 AGGAGAACCCCAAGATATCGAAGATGAGCAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 487 AGGAGAACCCCAAGATATCGAAGATGAGCAAGAAACAAATTCAGTCCCTTCTCCATG 546  
Qy 481 AAATGGTACACGCAAGACATGTTGAGGAGAAAGAACTTGGCAACAGAGATGAGCCACAG 540  
Db 547 AAATGGTACACGCAAGACATGTTGAGGAGAAAGAACTTGGCAACAGAGATGAGCCACAG 606  
Qy 541 GAGAAACACAAACAGAGAGGATGATGAGTTCTTATGCGGACTGATGATGATGATGATGATG 600  
Db 607 GAGAAACACAAACAGAGAGGATGATGAGTTCTTATGCGGACTGATGATGATGATGATGATG 666  
Qy 601 AGACCCCTGGAACTCGAAGATATCTCATGAAGAACCGAGCATAGTTACACAGTGAAGAGA 660  
Db 667 AGACCCCTGGAACTCGAAGATATCTCATGAAGAACCGAGCATAGTTACACAGTGAAGAGA 726  
Qy 661 CAGTTTCCAAAGACTGTAAATCAGGATATGGAAGAGATGATGCTGTGAGCAGGAAATCCAG 720  
Db 727 CAGTTTCCAAAGACTGTAAATCAGGATATGGAAGAGATGATGCTGTGAGCAGGAAATCCAG 786  
Qy 721 ATTCAGGTGAACCAAGTAGTAGAAGATGAAGATTCACACCATGATGATGATGATGATGAT 780  
Db 787 ATTCAGGTGAACCAAGTAGTAGAAGATGAAGATTCACACCATGATGATGATGATGATGAT 846  
Qy 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACTCTAGAAATGAAGGATAGAAATCA 840



```
Db      847  ACCAAGTCTATGAGGAACAAGCAGTATATGAAACCTCTAGAAAATGAAAGGATAGAAAATCA 906
Qy      841  CAGAAAGTAACTGCTCCCTCGAGGATAATCTGTGTAAGAAATCAAGGTAAATGTGTAAG 900
Db      907  CAGAAGTAACCTGCTCCCTCGAGGATAATCTGTGTAAGAAATCAAGGTAAATGTGTAAG 966
Qy      901  AAGTAAGCATTTTCTGTGTGGAAGAACAGCAGGAAAGTACACACAGAAAACAAATAGAAAA 960
Db      967  AAGTAAGCATTTTCTGTGTGGAAGAACAGCAGGAAAGTACACACAGAAAACAAATAGAAAA 1026
Qy      961  CAGATGATCCAGAACAAAAGCAAAAGTTTAAAGAAAAGAACCTTAACCTTTTAAATAAT 1020
Db      1027  CAGATGATCCAGAACAAAAGCAAAAGTTTAAAGAAAAGAACCTTAACCTTTTAAATAAT 1086
Qy      1021  TTGATAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCCGTAAAGGGGAAAA 1080
Db      1087  TTGATAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCCGTAAAGGGGAAAA 1146
Qy      1081  TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAACCAATACCTTCAGAGTCCACGAG 1140
Db      1147  TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAACCAATACCTTCAGAGTCCACGAG 1206
Qy      1141  CAAGATATGGGAAGGGCGAGTGTGAGGATGATTGGCTGAGAAGAGGAGAAGTAATGAGG 1200
Db      1207  CAAGATATGGGAAGGGCGAGTGTGAGGATGATTGGCTGAGAAGAGGAGAAGTAATGAGG 1266
Qy      1201  TGCTACCTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGCTCCGCGAG 1260
Db      1267  TGCTACCTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGCTCCGCGAG 1326
Qy      1261  ACCTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGCAACAATTTCTAGGTCAATGCA 1320
Db      1327  ACCTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGCAACAATTTCTAGGTCAATGCA 1386
Qy      1321  GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTTAA 1380
Db      1387  GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTTAA 1446
Qy      1381  AAAATGACCTTGGCTGGATACCTCTTGATAGAGATAATGCAATGCAAAAGAAAGTTT 1440
Db      1447  AAAATGACCTTGGCTGGATACCTCTTGATAGAGATAATGCAATGCAAAAGAAAGTTT 1506
Qy      1441  ATGAAGAGTCTCAGTGTGACACTATGATGCTTGTGAAGTCCATTTATGCTTCA 1500
Db      1507  ATGAAGAGTCTCAGTGTGACACTATGATGCTTGTGAAGTCCATTTATGCTTCA 1566
Qy      1501  TCCTGAAGGCACAGAACAAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAT 1560
Db      1567  TCCTGAAGGCACAGAACAAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAT 1626
Qy      1561  CCGGAGATCTCGGCATGATGGAGATTTTATTTCCACCTGGGGGATGCATGCAGAG 1620
Db      1627  CCGGAGATCTCGGCATGATGGAGATTTTATTTCCACCTGGGGGATGCATGCAGAG 1686
Qy      1621  GGGTTGGGAACAAAGAGGCATATAGTGTATGAGCTTGGGCAAGAGAGACACTTTG 1680
Db      1687  GGGTTGGGAACAAAGAGGCATATAGTGTATGAGCTTGGGCAAGAGAGACACTTTG 1746
Qy      1681  CATCTGTGGCAACCTCACTCAATATGATGACTGAAGCAACAGCCTTGGTGGG 1740
Db      1747  CATCTGTGGCAACCTCACTCAATATGATGACTGAAGCAACAGCCTTGGTGGG 1806
Qy      1741  CCCCAGAAAGAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTCGAAGTTAA 1800
Db      1807  CCCCAGAAAGAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTCGAAGTTAA 1866
Qy      1801  TCCGAGATGAAGCCCTTGAGTGAATGAATGAATGACTGAAGCAACAGCCTGAGGATG 1860
Db      1867  TCCGAGATGAAGCCCTTGAGTGAATGAATGAATGACTGAAGCAACAGCCTGAGGATG 1926
Qy      1861  AAAACCTGAGGAAAAGGGGACTGGACCCAGTTCACGCTGTGGCAGCAAGGAAGAA 1920
```

```
Db      1927  AAAACCTGAGGGAAAAAGGGGACTGGAGCCAGTTTCACTGTGTCACAAAGGAAGAA 1986
Qy      1921  ATGAAAAATGCTCAAGAGGAGCTCTTAAAACTGTGTACCTTACTAGAAAAAGTTCCCGGAG 1980
Db      1987  ATGAAAAATGCTCAAGAGGAGCTCTTAAAACTGTGTACCTTACTAGAAAAAGTTCCCGGAG 2046
Qy      1981  CAAACAGATGTCAGAAAGAGACAGATCAAAATATTTCCATCATGCACTCCCGGACTCACGTGT 2040
Db      2047  CAAACAGATGTCAGAAAGAGACAGATCAAAATATTTCCATCATGCACTCCCGGACTCACGTGT 2106
Qy      2041  GCGCGCACACAGGGCCCAAAATGCACTCCGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100
Db      2107  GCGCGCACACAGGGCCCAAAATGCACTCCGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2166
Qy      2101  AGGAAGCTGCAAGATTCGATGTGCCAACAGAGACAGGACCTGGAGGAAGCAAGGTGC 2160
Db      2167  AGGAAGCTGCAAGATTCGATGTGCCAACAGAGACAGGACCTGGAGGAAGCAAGGTGC 2226
Qy      2161  TCATCTTTGATGACTCTCTTTGAGCACAGAGTATGCGAGATGCTCATCTTTCCGGCTGA 2220
Db      2227  TCATCTTTGATGACTCTCTTTGAGCACAGAGTATGCGAGATGCTCATCTTTCCGGCTGA 2286
Qy      2221  TATTTCATCGTGGATGTGTGGCATCCGAACTCCGAACTGACACACAGCAGAGACGAGCTTCCAG 2280
Db      2287  TATTTCATCGTGGATGTGTGGCATCCGAACTCCGAACTGACACACAGCAGAGACGAGCTTCCAG 2346
Qy      2281  CAATTAGCATGAATTCATGCAAGCTTTGGAAACTCTGGAGAGA 2324
Db      2347  CAATTAGCATGAATTCATGCAAGCTTTGGAAACTCTGGAGAGA 2390
```

```
RESULT 4
HSU03109
LOCUS      2449 bp mRNA linear PRI 30-NOV-1995
DEFINITION Human aspartyl beta-hydroxylase mRNA, complete cds.
ACCESSION U03109
VERSION    U03109.1 GI:458031
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2449)
            Koriath, F., Gieffers, C. and Frey, J.
            Cloning and characterization of the human gene encoding aspartyl
            beta-hydroxylase
            Gene 150 (2), 395-399 (1994)
JOURNAL    MEDLINE 95121937
PUBMED     7821814
REFERENCE  2 (bases 1 to 2449)
            Koriath, F.
            Direct Submission
            Submitted (03-NOV-1993) Koriath, F., Fakultät fuer Chemie-Biochemie
            II, Universitaet Bielefeld, Universitaetsstrasse 25, Bielefeld,
            33615, Germany
FEATURES   Location/Qualifiers
            1..2449
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="As-5"
               /cell_line="mg63"
               /cell_type="osteosarcoma"
               /clone_lib="MG63-ZAP"
            1..77
               /function="hydroxylation of aspartyl and asparaginy
               residues"
               /codon_start=1
               /product="aspartyl beta-hydroxylase"
               /protein_id="AAA82108.1"
               /db_xref="GI:458032"
               /translation="MAQRKNAKSSGSSSGSGSTAGSSSPGARRETKHGHHG
```





Db 1009 ACCGAGACTATGATGAACAA----- 1028  
QY 841 CAGAAGTAACTGCTCCCTCGAGGATTAATCTGTAGAGANTTCAAGGTTAATTGTAGAAG 900  
Db 1029 -----GATCATGCTGTAGATAATTCAAAACACAACTTTTGAAG 1065  
QY 901 AAGTAAGCATTTTTCCTCTGTGAGAGACAGCAGGAGTACACACAGAAACAAATAGAAAAA 960  
Db 1066 AACCGCAATGCCCCCTGCGAGAGAAACAGCAAGAGTACACACAGAAACAAATAGAAAG 1125  
QY 961 CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAAAGAGCCCTAAACTTTTAAATAAT 1020  
Db 1126 CAGATGAGCCAGGAAAGAAAGGAAAGTGAAGAAAGAGAGCCCTAAACTTAAATAAT 1185  
QY 1021 TTGATAAGCATTTAAAGCTGAACTTGATGCTGAGAAAAAATCTCCGTAAAGGGGAAAAA 1080  
Db 1186 TTGATAAGCATTTAAAGCTGAGCTGATGCTGAGAAAAAATCTCCGTAAAGGGGAAAAA 1245  
QY 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACCAAAATACCTCAGAGTCCACGAG 1140  
Db 1246 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACCAAAATACCTCAGAGTCCACGAG 1305  
QY 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAAATGAGG 1200  
Db 1306 CTCGGTATGGGAAGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGCAATGAGA 1365  
QY 1201 TGCTAGTGGAGCCATCGAGACCTACCMAGAGGTGGCCAGGCTACCTGATGTCCTGCAG 1260  
Db 1366 TCCTGCGCAGGCGCATCGAGACCTACCAAGAGGCGGCGAGTCTGCTGATGCCCCACAG 1425  
QY 1261 ACCTGCTCAAGCTGATTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGCTCATATGA 1320  
Db 1426 ACTTGTGAAGCTGAGCTGGAAGCGAGGTGAGACAGACAGCAGTTTCTAGTTCACATGA 1485  
QY 1321 GAGGTTCCTCTGTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACCTTCTTAA 1380  
Db 1486 GAGGATCCTTCTTACTCTCAGAAATTTAGTTCAACTATTTCTCTGATGATACTGCTTAA 1545  
QY 1381 AAATGACCTTGGCTGGGATACCTCTGTAGAGGATTAATGCAATGCAAGAAAGTTT 1440  
Db 1546 AGAATGACCTTGGAGTTGGATACCTCTGTAGAGTGAACAATGACAGTGCCCAAGAGTCT 1605  
QY 1441 ATGAAGAGTGTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
Db 1606 ATGAAGAGTGTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTATGGCTTCA 1665  
QY 1501 TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGAAATAGAA 1560  
Db 1666 TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCTATCTGAAGGAAGAAATAGAA 1725  
QY 1561 CCGGAGATCTGGCCTGATGATGGAGATTTTATTCACCTGGGGGATGCCATGCAGA 1620  
Db 1726 CTGGGGATCCTGGTACTGATGATGGAGATTTTATTTTACCTGGGGAGCGCATGCAGA 1785  
QY 1621 GGGTTGGGAACAAAGGCAATATAAGTGTATGATGCTTTGGGCACAGAGAGACACTTTG 1680  
Db 1786 GAGTTGGAAACAAAGAGGCATATCGTGTGTACAGCTGGGCACACAGAGGGGCCACTTTG 1845  
QY 1681 CATCTGTCTGGCAACGCTCACTCTACATGTGAATGGAGTGAAGAACACAGCCCTTGGTGA 1740  
Db 1846 CCTCGTGTGGCAGCGCTCACTGTACAACGTGAGGATTTGAAGCCAGCGCGTGGTGA 1905  
QY 1741 CCCCAGAAAGAGCGGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db 1906 CCCCCAGAGAGACCGGCTACACCGAGTTAGTGAATCTTTAGAAAGAACTGGAAGTTAA 1965  
QY 1801 TCCGAGATGAAGCCCTTGCAGTGAATGAATGAAGCCAAAGGTCTTCTTCTGCTGAGGATG 1860  
Db 1966 TCAGAGATGAAGCCCTTGCAGCAATGATAGAACCCATGCGCCTCTTCTGCTGAGGATG 2025  
QY 1861 AAACCTGAGGGAAGGGGAGCTGAGGCGAGTTACGCTGTGGCAGCAAGCAAGAGAA 1920

Db 2026 AGAACCTGAGGAAAAAGGTGACTGTGAGTTCAGTTTACGCTGTGGCAGCAAGAGAAAAA 2085  
QY 1921 ATGAAATATGCTGCAAGAGAGCTCTCTAAATACCTGTACTTACTAGAAAAAGTTCCCGAGA 1980  
Db 2086 ATGAAATATGCTGTAAGAGAGCTCTTAAGACCTGTTCTTACTAGATAAAATCCCTGAGA 2145  
QY 1981 CAACAGGATGCAAGAGAGAGAGAGATCAATATTTCCATCATGACACCCGGGACTCAGGTG 2040  
Db 2146 CAACAGGATGCGAAGAGAGAGAGATCAATATTTCCATCATGACACCCGGGAAACCCAGGTG 2205  
QY 2041 GCGCCGACACAGGCGCCCAAACTGAGCTCCGAATGACCTGGGCTTGGTGTGATTTCCCA 2100  
Db 2206 GCGCGCACACAGGAGCCCGAACTGTAGCTCCGAATGACCTGGGCTTGGTGTGATTTCCCA 2265  
QY 2101 AGGAAGGCTGCAAGATTTGATGTGCAACAGACAGACAGCTGGAGAGAGCAAGGTGC 2160  
Db 2266 AGGAAGGCTGCAAGATTTGATGTGCAACAGACAGACAGCTGGAGAGAGCAAGGTGC 2325  
QY 2161 TCATCTTTGATGACTCTTTGAGCAGGAGTATGCGAGATGCCTCATCTTTCCGCTCA 2220  
Db 2326 TCATCTTTGACGACTCTTTTGAACACAGAGGTGTGGCAGGATGCCGCTCTTTCCGCTCA 2385  
QY 2221 TATTTCATCTGATGTGTGCGATCCGGAATCTGCAACAGACAGACAGCTTTCCAG 2280  
Db 2386 TCTTCATTTGATGTGTGCGACCTTGAGCTGACACCCCATCAGAGACGAGCTTCTCTG 2445  
QY 2281 CAATTTAGCATGAATTTATGCAAGTTTGGAAATCTTGAGAGA 2324  
Db 2446 CAATTTAGCAT---TCTGTGCGGCTCAGGACACGCTGGAGAGA 2486

RESULT 6  
AF289486 4419 bp mRNA linear ROD 18-DEC-2000  
LOCUS Mus musculus aspartyl beta-hydroxylase 4.5 kb transcript (Asph)  
DEFINITION mRNA, complete cds, alternatively spliced.  
ACCESSION AF289486  
VERSION AF289486.1 GI:11878109  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 4419)  
AUTHORS Dinchuk,J.E., Henderson,N.L., Burn,T.C., Huber,R., Ho,S.P., Link,J., O'Neil,K.T., Focht,R.J., Scully,M.S., Hollis,J.M., Hollis,G.F. and Friedman,P.A.  
TITLE Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junctin  
JOURNAL J. Biol. Chem. 275 (50), 39543-39554 (2000)  
MEDLINE 20564328  
PUBMED 10956665  
REFERENCE 2 (bases 1 to 4419)  
AUTHORS Dinchuk,J.E., Henderson,N.L., Burn,T.C., Hollis,G.F. and Friedman,P.A.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA  
FEATURES  
source  
1..4419  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/chromosome="8"  
/map="8q12 at STS WT-11767"  
/tissue\_type="liver"  
1..4419  
/gene="Asph"  
1..174  
/gene="Asph"  
175..2394

```

/gene="Asph"
/note="BAH; asparaginyl beta-hydroxylase; Asph;
alternatively spliced"
/codon_start=1
/product="aspartyl beta-hydroxylase 4.5 kb transcript"
/protein_id="AAG40808.1"
/db_xref="GI:11878110"
/translation="MAPRKNAGKGGGSSSGSGSGSGSPSTGSGSSSPGARRR
AKGHGKNRGGTSGGSPFTMWLALLGVWTSVAVWEDLVDYBEVLGKLVGDAD
GDGDFVDVDAKVLGLKERSPSERTPPEAEHAELEBQAPEDADLQNVDEVKQIQ
SLLQSVHTQDLEADLAGERQPEVEFLVITUSDDRFEDLEFPVHEEIDETVHE
TASQNHNDMEWMTNEQENSDFEAVTAGVLLPHAEVRHQDYDEPYFSEHEGV
RISQNTIDSSISIEELINVASVEEQDTPPVKKKKPKLLNKFKTKIABELDAEKLK
RGKIEAVNAPFELVKRYPOSPPARYQAQCEDDLAEKQSNELVRAIETYOEAADL
PDATDLVLSLKRSEKQFLGHRGSLTLQRLVQLPFSDDLKNDLGVGLLGD
YFHLGDAMQVGNKEAYKWEYELGKRGHFPASVQSRSLYNVGLKAPQWMTPRETYTE
LVKSLERNKLIIRDGLVMYDKAGLFLPEDENLRKGDMSQFTLMQOGRKNENACKG
APKTCALLEKFTTGCRRGOIKYSIMHPGTHVWPHPTNCLRWHLGLVTPKGGCK
IRCANETRTWEEGKVLIFDDSPHEVWDASSFRLIFIVDWHPELTPQRRSLPAI"
2395..4419
3'UTR
/gene="Asph"
polyA_signal
4379..4384
/gene="Asph"
ORIGIN

```

Query Match 62.4%; Score 1450.2; DB 10; Length 4419;  
 Best Local Similarity 79.5%; Pred. No. 1.1e-303;  
 Matches 1833; Conservative 0; Mismatches 378; Indels 96; Gaps 6;

```

QY 15 GCCCAGCGTAAAGATGCCAAGACGCGGCAACAGCAGCAGCAGCGGCTCCGCGACGGT 74
DB 193 GCCAAGGCGCGCGCGCAACAGCAGCAGCAGCGGCGCAGCGGCTCGGCGACGGT 252
QY 75 AGCACGAGTGGCGGACGACGCGTACGACGAGTACGCGCGGCGCTCGAGAGAGCAAG 122
DB 253 AGCCGAGCAGCGGTAGCAGCGTACGACGAGTACGCGCGGCGCTCGAGAGAGCAAG 312
QY 123 CATGAGGACACAGAGATGGAGAGAGCGGACTCTCGGAGACTTCATCTTCACGTGG 182
DB 313 CACGAGGACACAGAGATGGAGAGAGCGGATTCAGAGAGGTCCTTTTCACATGG 372
QY 183 TTTATGTGATTCGATGCTGGCGGTCTGGAACATCTGATGCTGCTGTTGGTTGATCTT 242
DB 373 TTCATGTGATTCGATGCTGCGCGCTCTGGAACATCTGCGCTGCTGCTGCTGCTGCTG 432
QY 243 GTTGACTATGAGGAAGTTCTAGGAAACTAGGAACTATGATGCTGATGCTGATGGAGAT 302
DB 433 GTCGATTTATGAAGAAGTTCTAGGAAACTAGGAGTCTATGATCGCGATCGCATGGAC 492
QY 303 TTTGATGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGCA 362
DB 493 TTTGATGTGATGATGCCAAAGTTTATTAGGCTTTAAAGAAAGATCTCTCTGAGAGG 552
QY 363 GCAGTCCGCCAAGAGCGCTGAGCCACACACTAGCGCCGAGGAGAGGTTCTGTGGAG 422
DB 553 ACATTTCCACCC---GGAGCAGAGACTCACGCTGAGCTGAGGAGAGCGCCCTGAGGG 609
QY 423 GCAGACCCCGAGAAATATCGAGATGAAGCAAAAGAAACAATTCAGTCCCTTCTCATGAA 482
DB 610 GCAGACATCCAGAAATGTTGAAGATGAAGTAAAGAAACAATTCAGTCCCTTCTCATGAA 669
QY 483 ATGTCACACGAGACATGTTGAGGAGAGACTTGCACAAAGAGAGATGGACCCACAGGA 542
DB 670 TCAGTACACAGACCATG-----ACCTTGAAGCAGATGGCGTGGCAGGA 714
QY 543 GAACCAACACAGAGATGATGAGTTTCTTATGGGACTGATGATGATGATGATGATGATG 602
DB 715 GAACCAACAGCGGATGAGGACTTCTCACAGTACCGCAGCAGTGAAGATGATGATGATG 774
QY 603 ACCCTGGAACTTGAAAGTATCTCATGAAGAAACCGAGCATAGTTTACCAGCTGGAAAGACA 662
DB 775 GACCTGGAAACCCGGGACAGTTTCATGAAGAAATCGAGGATACCTTACATGTCGAAGACA 834

```

```

QY 663 GTTTCACAGACTGTAACTCAGGATATGGAAGAGATGATGTCTGAGCAGGAAAAATCCAGAT 722
DB 835 GCATCGCAGAACCCATCCAAATGACATGGAAGAGATGACGAATGACGAGGAAAAATTCAGAT 894
QY 723 TCCAGTGAACCAAGTAGTAGAAGATGAAGATTTGCCACCATGATGATGATGATGATGATGATG 782
DB 895 CCCAGTGAAGCAGTGCAGATGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 954
QY 783 CAAAGTCTATGAGGAAACAAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAAATCACA 842
DB 955 CAAAGACTATGATGAAC---CAGTATATGAACCTTCAGAGCAGCAAGGGGTCGAGATTTCA 1011
QY 843 GAAGTAACTGCTCCCTCGAGGATATCTCTAGAAATTCACAGGTAAATCTCAGAGAA 902
DB 1012 -----GATTAACCATGATGATTCAGAGCAATATCTCAGAGAA 1050
QY 903 GTAAGCAATTTTCTCTGGAAGAACAGCAGGAAGTACACACAGAAAACAAATAGAAAAACA 962
DB 1051 ATAATGTCCGCTCTGTGTAAGAACAGCAGACACACACC----- 1091
QY 963 GATGATCCAGAACAAAAAGCTTAAGAAAAAGAGCCTTAAACCTTTTAAATAAATTT 1022
DB 1092 -----AGTTAAGAAAGAAAGCCTTAAACCTTCTGAACAAATTT 1128
QY 1023 GATAAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAACTCCGTAAAGGGGAAAAATTT 1082
DB 1129 GATAAGCAATTAAGGCTGAGCTGATGCTGCAGAAAAGCTCCGAAAAGGGGTAAATTT 1188
QY 1083 GAGCAGCAGTGAATGCATTTAAAGAACTAGTAGCMAATACCTCTCAGAGTCCAGAGCA 1142
DB 1189 GAGAGCAGTGAATGCATTTGAAGAACTGGTTGGAAGTACCTCTCAGAGCCACAGCA 1248
QY 1143 AGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGATGAGGTG 1202
DB 1249 AGATATGGCAAGCGCAGTGTGGAAGATGACTTTGGCAGAGAAAGCAGAGCAATGAGGTT 1308
QY 1203 CTACGTGGAGCCATCGAGACCTTACCAAGAGTGGCAGGCTACCTGATGTCTCCCTGACAG 1262
DB 1309 CTGCGAGGCGCATCGAGACCTTACAGAGGCGGCGACCTGCTGTGATGCTCCCTACAGAC 1368
QY 1263 CTGCTCAAGCTCAGTTTCAAGCGTGCCTCAGACAGGCAAAATTTCTAGGTCATATGAGA 1322
DB 1369 CTGCTCAAGCTCAGTTTCAAGCGAGGTTCGAAACGCGACGATTTCTTAGGTCATATGAGA 1428
QY 1323 GGTTCCTGCTTACCTCCAGAGATTAAGTTCAACTATTTCCCAATGATCTTCTCTTAAAA 1382
DB 1429 GGTTCCTTACTTACCCTCCAGAGACTAGTTTCACTGCTTCCCTAGTGTACTTACCTTAAA 1488
QY 1383 AATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAATGCAAGAAAGAGTTTAT 1442
DB 1489 AACGACCTTGGCGTAGGATACCTCTTGTGGGAGACAACAGCAGTGCACCAAGAGGTTTAC 1548
QY 1443 GAAGAGTCTCAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTCATC 1502
DB 1549 GAAGAGTCTTAAATGTGACACCAATGATGGCTTCGCTTAAAGTGCATTCAGGCTTCATC 1608
QY 1503 CTGGAAGCAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGAAAGGAAATGAAATCC 1562
DB 1609 CTGGAAGCAGAACAAAGATATCTGAGAGCATTTCCCTACTTAAAGAAAGGAAATCGAATCT 1668
QY 1563 GAGATCTCTGGACGTGATGATGAGGATTTTATTTCCACCTGGGGGATGCCATGCGAGG 1622
DB 1669 GGGGACCTTGGCAGCGATGATGTCGCGGTTTATCTTCCACTTGGGGGATGCCATGCGAGG 1728
QY 1623 GTTGGGAAACAAAGAGGATATAAGTGTATGAGCTTGGGCAACAGAGAGGACACTTTGCA 1682
DB 1729 GTCGGAAACAAAGAGGATATAAGTGTATGAACTTGGGCAACAGAGAGGACACTTTTGC 1788
QY 1683 TCTGTCTGGCAACCTCCTCCTACATGTGAATGCACTGAAAGCAGACCTTGGTGACC 1742
DB 1789 TCTGTCTGGCAGCGTTCCCTCTACATGTGAATGCTGAAGGCTCAGCGCTGCTGGGACA 1848

```

Qy	1743	CCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACCTGGAGTTAATC	1802
Db	1849	CCAGGAGAGACTGGCTACACAGAGCTAGTGAAGTCTTTTAAAGAAACCTGGAGTTAATC	1908
Qy	1803	CGAGATGAAGCCCTTGAGTGATGAAGCAAGCAAGGCTCTCTCTGCTGAGGATGAA	1862
Db	1909	CGTGATGAAGCCCTTGAGTGATGAAGCAAGCAAGGCTCTCTCTGCTGAGGAGCA	1968
Qy	1863	AACCTGAGGGAAGAAAGGGGACTGGAGCTGAGCTTCAAGCTGGGAGCAAGAAAT	1922
Db	1969	AACCTTGGGAGAGGCGACTGGAGCTGAGCTTCAAGCTGGGAGCAAGAAAT	2028
Qy	1923	GAAATGCTGCAAGAGAGCTCTAAACCTGTACCTTACTAGAAAGTCCCGGAGACA	1982
Db	2029	GAGAATGCTGTAAGAGAGCGCTTAAGACCTGTGCTTTACTAGAAAGTTCCTCGAACA	2088
Qy	1983	ACAGGATGCAAGAGAGAGAGATCAATATATTCATCATGACACCCCGGACTCATCGTGTGG	2042
Db	2089	ACAGGATGCAAGAGAGAGAGATCAATATATTCATCATGACACCCCGGACTCATCGTGTGG	2148
Qy	2043	CCGCACACAGGGCCCAACAACTGAGGCTCGAATGCACTGGGCTTGGTATCCCAAG	2102
Db	2149	CCGCATACAGGACCCACAACTGAGGCTCGAATGCACTGGGCTTGGTATCCCAAG	2208
Qy	2103	GAGGCTGCAAGATTCGATGTGCCAAGAGACAGGACCTGGGAGGAGGCAAGTGTCTC	2162
Db	2209	GAGGCTGCAAGATTCGATGTGCCAAGAGAGACAGGACCTGGGAGGAGGCAAGTGTCTC	2268
Qy	2163	ATCTTTGATGACTCTCTTTGAGCAGCAGGATGAGGAGGAGGCTCATCTTTCGGCTGATA	2222
Db	2269	ATCTTTGATGACTCTCTTTGAGCAGCAGGATGAGGAGGAGGCTCATCTTTCGGCTGATA	2328
Qy	2223	TTTCATGCTGGATGTGGTATCGGAACTGACACACAGAGAGAGGAGGCTTCCAGCA	2282
Db	2329	TTTCATGCTGGATGTGGTATCGGAACTGACACACAGAGAGAGGAGGCTTCCAGCA	2388
Qy	2283	ATTAGCATGAATTCATGCAAGCTTGG	2309
Db	2389	ATTGAAAGGCACTGACGAGACTTGG	2415
RESULT 7			
AF289487		6629 bp mRNA linear ROD 18-DEC-2000	
LOCUS		Mus musculus aspartyl beta-hydroxylase 6.6 kb transcript (Asph)	
DEFINITION		mRNA, complete cds; alternatively spliced.	
ACCESSION		AF289487	
VERSION		AF289487.1 GI:11878111	
KEYWORDS		Mus musculus (house mouse)	
SOURCE		Mus musculus	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE		1 (bases 1 to 6629)	
AUTHORS		Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P., Link, J., O'Neil, K.T., Focht, R.J., Scully, M.S., Hollis, J.M., Hollis, G.F. and Friedman, P.A.	
TITLE		Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with Junctin	
JOURNAL		J. Biol. Chem. 275 (50), 39543-39554 (2000)	
MEDLINE		20564328	
PUBMED		10956665	
REFERENCE		2 (bases 1 to 6629)	
AUTHORS		Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and Friedman, P.A.	
TITLE		Direct Submission	
JOURNAL		Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA	
FEATURES		Location/Qualifiers	
source		1..6629	
		/organism="Mus musculus"	
		/mol_type="mRNA"	

gene	/strain="BALB/c"	
5'UTR	/db_xref="taxon:10090"	
	/chromosome="8"	
	/map="8q12 at STS WT-11767"	
	/tissue_type="liver"	
	1..6629	
	/gene="Asph"	
	1..174	
	/gene="Asph"	
	175..2394	
	/gene="Asph"	
	/note="BAH; asparaginyl beta-hydroxylase; Asph; alternatively spliced"	
	/codon_start=1	
	/product="aspartyl beta-hydroxylase 6.6 kb transcript"	
	/protein_id="AAG40809.1"	
	/db_xref="GI:11878112"	
	/translation="MAPRKNAGGGSSSGSGSGSGSPSTGSSGSSPGARRE AKHGKNGRRGGISGGSFTWPMVIALLGWTSVAVVWFDLVDYEVLLGKLVYDAD GDGDFDADAKVLLGLKERSPSERTFPPEAEHAELEBQAPEGADIQNVDEVEKEIQI SLLQESVHTDHLDEADGLAGEPQPEVEDFLTVTSDDRFEDLEPFGTVEHEIEDYHVE FTSDQHPNMEENTQENSDESEAVTDGAVLLPHAEVRHQDYDEPVYEPSEHEGV RKTLEAVNAPEELVRKYPOSPRARYKAQCEDDLAEKQRSNEVLRRAIETVQEAADL PDAITDLKLSLKRSEKQQLGHMRGSLTLQRLVOLFPSDTTLKNDLGVGLLLGD NDSAKKVVEEVLNVTNDGFAKHGYPILKAQNKLSIESIPVYKEGIESGDPGTTDGRF YFHLGDAMORVGNKEAYKYWELGHRGFASFVWQKSLNVNGLKAQPMWTPRETQYTE LKVSLEKNWKLIRDEGLMVDKQIKYIMPHGTWHPHTGTNCRMLHLGLVLPKRGCK LKPTCALLKPFSETTCGRGQIKYIMPHGTWHPHTGTNCRMLHLGLVLPKRGCK TSCANETRTWEEGKVLIPDSSFEHEVWQDASSFLRIFIVDVWHPHLPQTQRRSLPAI"	
	2395..6629	
3'UTR	/gene="Asph"	
	4379..4384	
	/gene="Asph"	
	6573..6578	
	/gene="Asph"	
ORIGIN		
	Query Match	62.4%; Score 1450.2; DB 10; Length 6629;
	Best Local Similarity	79.5%; Pred. No. 1.1e-303;
	Matches 1833; Conservative	0; Mismatches 378; Indels 96; Gaps 6;
QY	15	GCCAGCGTAAGAATCCAAAGACGACGCGCAACAGCAGCAGCAGCGCGCTCCGCGACGGT 74
DB	193	GCCAAAGGCGCGCGCGCAACAGCAGCAGCAGCGCGCGCGCGCTCCGCGACGGT 252
QY	75	AGCAGCAGTCCGCGCAGCAGC-----AGCCCCGGGCCCCGAGAGACAAAG 122
DB	253	AGCCCGAGCAGCGGTAGCAGCGGTAGCAGCAGTAGCCCCGGGCTCGGAGAGAGCAAG 312
QY	123	CATGAGGACACAGATGGAGGAAGCGGACTCTCGGGAACCTTCATCTTCACGTGG 182
DB	313	CACGAGGACACAGATGGAGGAAGAGAGGAGATTCAGAGGGTCTCTTTTCATGCG 372
QY	183	TTTATGTGATTGTCATTGCTGGCGCTCTGGACATCTGTAGCTGTGCTTGTGTTGATCTT 242
DB	373	TTTATGTGATTGTCATTGCTGGCGCTCTGGACATCTGTGCTGTGCTGTTGATCTT 432
QY	243	GTTCACATGAGGAAGTTCAGGAAACTAGGAACTATGATGCTGATGCTGATGATGAGAT 302
DB	433	GTCGATTATGAAGAAGTTCAGGAAACTAGGAGTCTATGATGCGGATGCGGATGAGAC 492
QY	303	TTTGTATGTGATGATGCCAAAGTTTATAGGACTTAAAGAGAGATCTACTTTCAGACCA 362
DB	493	TTTGTATGTGATGATGCCAAAGTTTATAGGCTTAAAGAAAGATCTCTCTCTGAGAG 552
QY	363	GCAGTCCCGCCAGAGAGGCTGAGCCACACTGAGCCCGAGGAGGAGTTCCTGTGGAG 422
DB	553	ACATTTCCACC---GGAGGCAGAGACTCACGCTGAGCTGGAGGAGCGCCCTGAGGG 609
QY	423	GCAGAACCCCGAGATATCGAAGATGAACAAAGAAACAAATTCAGTCCCTTCTCCATGAA 482
DB	610	GCAGACATCCAGAAATGTTGAAGATGAAGTAAGGAACAAATTCAGTCCCTTCTTCAGGA 669

QY 483 ATGTTACACGACGACATGTTGAGGGAGAGAGACTTGGCAACAAGAGATGGACCCACAGA 542  
Db 670 TCAGTACACAGACACATG-----ACCTTGAGCAGATGGCGTGGCAGGA 714  
QY 543 GAAACACACAGAGAGATGATGATTTCTATGGGACCTGATGTAGATGATGATTTGAG 602  
Db 715 GAAACACAGCGGAGGTTGAGGACCTTCTCACAGTACCGGACGAGTATGATGATTTGAG 774  
QY 603 ACCCTGGAACCTGAAATGATCTCATGAAGAAACCGAGCATAGTTACACAGTGGAGAGACA 662  
Db 775 GACCTGGAACCGGAGCAGTTTATGAAGAAATTCAGGATACCTTACCATGTGGAAGACACA 834  
QY 663 GTTTTACACAGACTGTAATCAGGATATGGAAGATGATGTCTGAGCAGAGAAATCCAGAT 722  
Db 835 GCATCGCAGAACCATCCAATGCATGGAAGATGACGATGAGCAGAGAAATTCAGAT 894  
QY 723 TCCAGTCAACAGCTAGTAGAAGATGAAGATTTGCACCATGATGATGATGATGATGATGAT 782  
Db 895 CCCAGTGAAGCAGTACAGATGAGTGTGCTGCTGCCCATGACAGAGAGATGATGATGATGAT 954  
QY 783 CAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAAGGATGAGAAATCA 842  
Db 955 CAAGACTATGATGAAC---CAGTATATGAACCTTTCAGAGCAGCAAGAGGCTGAGATTTCA 1011  
QY 843 GAAGTAATCTCTCCCTCAGGATATCTCTAGAAAGATTCACAGGTAAATTTGAGAGAA 902  
Db 1012 -----GATAACACCATAGATGATTCAGACATAATCTCAGAGAA 1050  
QY 903 GTAAGCATTTTCTCTGGAAGAACAGCAGGAGATGATCCACCCAGAGAAATAGAAACA 962  
Db 1051 ATAAATGTCCTCTCTGGAAGAACAGCAGACACACACAC----- 1091  
QY 963 GATGATCCAGAACAAAGAAAGTTAAAGAAAGAGCTTAACTTTTAAATAATTT 1022  
Db 1092 -----AGTTAAGAAAGAGAGAGCTTAACTTTCTGAACAAATTT 1128  
QY 1023 GATAAGACTATAAAGCTGAACCTTGATGCTGCAGAAAGAACTCCGTAAAGGGGAGAAATTT 1082  
Db 1129 GATAAGAACTAATTAAGCTGAGCTGATGATGCTGCAGAAAGAGCTCCGGAAGAGGGGTAAATTT 1188  
QY 1083 GAGGAACAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAGCA 1142  
Db 1189 GAGGAACAGTGAATGCAATTTGAAGAACTGGTTCGAAGTACCTCAGAGGCCACGAGCA 1248  
QY 1143 AGATATGGGAGCGGCGAGTCTGAGGATGATTTGGCTGAGAGAGGAGAGATTAATGAGTG 1202  
Db 1249 AGATATGGCAAGCGGCGAGTGTGAAGATGATTTGGCAGAGAGGAGAGAGAGATGAGGTT 1308  
QY 1203 CTAGTGGAGCCATCGAGACCTTACCAAGAGTGGCCAGCCTTACCTGATGTCCCTCGACAC 1262  
Db 1309 CTGGCAGGGCCATCGAGACCTTACCAAGAGGCGAGCCGACCTGCTGATGCCCTACAGAC 1368  
QY 1263 CTGCTGAAGCTGAGTTGAGGCTCGCTCAGACAGGCAACAAATTTCTAGTTCATATGAGA 1322  
Db 1369 CTGCTGAAGCTGAGTTGAGGCGAAGGTCGGAACGGCAGCAGTTTCTAGTTCATATGAGA 1428  
QY 1323 GGTTCCTGCTTACCTGCTGAGAGATTTAGTTCAACTATTTCCCAATGATACTTCTTTAAA 1382  
Db 1429 GGTTCCTGCTTACCTGCTGAGAGATTTAGTTCAACTATTTCCCAATGATACTTCTTTAAA 1488  
QY 1383 AATGACCTTGGCTGGGATACCTCTTGATAGGAGATTAATGACAAATGCAAGAAAGTTAT 1442  
Db 1489 AACGACCTTGGCTGGGATACCTCTTGATAGGAGATTAATGACAAATGCAAGAAAGTTAT 1548  
QY 1443 GAAGAGTCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATATGCTTCTATC 1502  
Db 1549 GAAGAGTGTAAATGTGACACCAATGATGGCTTCGCTAAAGTGCATTTACGGCTTCTATC 1608  
QY 1503 CTGAAGGCACAGAAATAATTCCTGAGAGCATCCATATTTTAAAGGAAGAAATAGAAATCC 1562  
Db 1609 CTGAAGGCACAGAAATAATTCCTGAGAGCATTCCTTAAAGGAAGAAATGCAATCT 1668

QY 1563 CGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCCAGAG 1622  
Db 1669 GGGGACCTTGGCAGGATGATGCGCGTTTACTTTCCACTTGGGGGATGCCATGCCAGAG 1728  
QY 1623 GTTGGGCAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGAGACACTTTGCA 1682  
Db 1729 GTGCGGCAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGAGACACTTTGCA 1788  
QY 1683 TCTGTCTGGCAACGCTCATCTACATATGTGAATGAGTGAAGCAGACAGCCTTGGTGGACC 1742  
Db 1789 TCTGTCTGGCAGCGTTCCCTCTACAATGTGAATGTCTGAAGGCTCAGCGTGGTGACA 1848  
QY 1743 CCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATC 1802  
Db 1849 CCCAGGAGAGCTGGCTACACAGAGCTAGTGAAGTCTTTAGAGAGAACTGGAAGTTAATC 1908  
QY 1803 CGAGATGAAGCGCTTGCAGTATGATGAAGCAAGAGTCTTCTCTGCTGAGGATGAA 1862  
Db 1909 CGTATGAAGCGCTTGCAGTATGATGAAGCAAGAGTCTTCTCTGCTGAGGATGAA 1968  
QY 1863 AACCTGAGGAAACGGGCTGAGAGCCAGTTTCAAGTGTGGCAGCAAGAGAAAGAAAT 1922  
Db 1969 AACCTTGGGAGAGGCGGCTGAGAGCCAGTTTCAAGTGTGGCAGCAAGAGAAAGAAAT 2028  
QY 1923 GAAATGCTTGCAGAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGGAGACA 1982  
Db 2029 GAGAATGCTTGAAGAGCGCTTAAAGCTGTGCTTTTACTAGAAAAGTTTCCGAAACA 2088  
QY 1983 ACAGATGACAGAGAGAGAGAGATCAAAATTTCCATCATGACCCCGGAGCTACGTTGG 2042  
Db 2089 ACAGATGACAGAGAGAGAGAGATCAAAATTTCCATCATGACCCCGGAGCTACGTTGG 2148  
QY 2043 CCGCACACAGGCGCCACAACTGCGAGGCTTCCGAATGACCTTGGCTTGGTGTATCCCAAG 2102  
Db 2149 CCGCATACAGGCGCCACAACTGCGAGGCTTCCGAATGACCTTGGGTTTGGTGTATCCCAAG 2208  
QY 2103 GAAGGCTGCAAGATTCGATGTGCCAACGAGACGAGGACCTGGGAGGAGCAAGGTGCTC 2162  
Db 2209 GAAGGCTGCAAGATTCGATGTGCCAACGAGACGAGGACCTGGGAGGAGCAAGGTGCTC 2268  
QY 2163 ATCTTTGATGACCTTCTTCCAGCAGAGGTATGGGAGAGTCTCATCTTTCCGGCTGATA 2222  
Db 2269 ATCTTTGATGACCTTCTTCCAGCAGAGGTATGGGAGAGTCTCATCTTTCCGGCTGATA 2328  
QY 2223 TTGATCTGATGTGGCATCCGAACTGACACACAGCAGAGAGAGAGAGAGAGAGAGAGAG 2282  
Db 2329 TTGATCTGATGTGGCATCCGAGCTGACCCCGGAGCTGACCCCTCAGCAGAGAGAGAGAG 2388  
QY 2283 ATTTAGCATGAATTCATGCAAGCTTGG 2309  
Db 2389 ATTTGAAAGGCACTGACGAGACTTGG 2415

## RESULT 8

LOCUS CQ729666 1970 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 15600 from Patent WO02068579.  
ACCESSION CQ729666  
VERSION CQ729666.1 GI:42301377  
KEYWORDS

## SOURCE

ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof

## JOURNAL

Patent: WO 02068579-A 15600 06-SEP-2002;

## FEATURES

source

PE Corporation (NY) (US)  
Location/Qualifiers  
1..1970



Query Match		52.3%; Score 1215.8; DB 6; Length 1970;
Best Local Similarity		90.2%; Pred. No. 7.5e-253;
Matches 1299; Conservative		0; Mismatches 141; Indels 0; Gaps 0;
ORIGIN		
Qy	849	ACTGCTCCCTCCTGAGGATAATCCTGTAGAGATTTCACAGGTAAATGTAGAGAAAGTAAGC 908
Db	531	ACTGAGCCGAGGAGCAGGTTCTGTGTGAGCGAGAACCCAGAAATATCGAAGATGAAGCA 590
Qy	909	ATTTTTCCTGTGGAAGAACACGAGGAAGTACCCACGAGAAACAAATAGAAAAACAGATGAT 968
Db	591	AAAGAACAAATTCAGTCCCTTCCTCCATGAAATGGTACACGACAGAACATGTTTGAGGGAGAA 650
Qy	969	CCAGAAACAAAGCAAAAGTTAAGAAAAAGAGCCTAAACCTTTTAAATAATTTGATAAG 1028
Db	651	GACTTGCACAAAGAGATGGACCCACAGAGGAACCAACAAAGAGGATGATGAGTTTCTT 710
Qy	1029	ACTATTAAAGCTGAACCTTGATGCTGCAGAAAACTCCGTAAAGGGGAAAAATTTGAGGAA 1088
Db	711	ATGGCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 770
Qy	1089	GCAGTGAATGCATTTAAAGAACTAGTAGCGAAATPACCTCAGAGTCCACGAGCAAGATAT 1148
Db	771	GCAGTGAATGCATTTAAAGAACTAGTAGCGAAATPACCTCAGAGTCCACGAGCAAGATAT 830
Qy	1149	GGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGTAATGAGTGTCTACGT 1208
Db	831	GGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGTAATGAGTGTCTACGT 890
Qy	1209	GGACCCATCGAGACCTACCAAGAGTGCAGAGCTCCTGATGTCCTGCAGACCTCCTG 1268
Db	891	GGACCCATCGAGACCTACCAAGAGTGCAGAGCTCCTGATGTCCTGCAGACCTCCTG 950
Qy	1269	AAGCTGATGTTGAAGCGTCGTACAGAGGCAACAAATTTCTAGTGCATATGAGAGGTTCC 1328
Db	951	AAGCTGATGTTGAAGCGTCGTACAGAGGCAACAAATTTCTAGTGCATATGAGAGGTTCC 1010
Qy	1329	CTGCTTACCTGCGAGAGATGATTTCAACTATTTCCCAATGATCTCTTAAATAATGAC 1388
Db	1011	CTGCTTACCTGCGAGAGATGATTTCAACTATTTCCCAATGATCTCTTAAATAATGAC 1070
Qy	1389	CTTGGCGTGGATACCTCTTCATAGGAGATTAATCAATGCAAGAAAGTTTATGAAGAG 1448
Db	1071	CTTGGCGTGGATACCTCTTCATAGGAGATTAATCAATGCAAGAAAGTTTATGAAGAG 1130
Qy	1449	GTGCTGATGTGACACCTAATGATGGCTTTGCTAAAGTCCATTAATGCTTCATCCTGAAG 1508
Db	1131	GTGCTGATGTGACACCTAATGATGGCTTTGCTAAAGTCCATTAATGCTTCATCCTGAAG 1190
Qy	1509	GCACAGAACAAATTTGCTGAGAGATCCCATATTTAAAGAGAGAAATAGATCCGGAGAT 1568
Db	1191	GCACAGAACAAATTTGCTGAGAGATCCCATATTTAAAGAGAGAAATAGATCCGGAGAT 1250
Qy	1569	CCTGCGCTGATGATGGAGATTTTATTTCCACTGGGGATGCGATGCGAGAGGTTGGG 1628
Db	1251	CCTGCGCTGATGATGGAGATTTTATTTCCACTGGGGATGCGATGCGAGAGGTTGGG 1310
Qy	1629	AACAAAGGCGATATAAGTGGTATGAGCTTGGGCACAGAGAGACACTTTGCAATCTGTC 1688
Db	1311	AACAAAGGCGATATAAGTGGTATGAGCTTGGGCACAGAGAGACACTTTGCAATCTGTC 1370
Qy	1689	TGGCAACGCTCACTCTACAAATGTAATGGACTGAAGCAGACGCTTGGTGCACCCCAAAA 1748
Db	1371	TGGCAACGCTCACTCTACAAATGTAATGGACTGAAGCAGACGCTTGGTGCACCCCAAAA 1430
Qy	1749	GAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGAT 1808
Db	1431	GAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGAT 1490
Qy	1809	GAAGGCTTGCAGTGTGATGATTAAGCAAAAGGTCTCTTCTGCTGAGGATGAACACCTG 1868
Db	1491	GAAGGCTTGCAGTGTGATTAAGCAAAAGGTCTCTTCTGCTGAGGATGAACACCTG 1550
Qy	1869	AGGAAAAAGGGGACTGGAGCCAGTTTCAGCTCTGCGAGCAAGAAAGAAAGAAATGAAAAAT 1928
Db	1551	AGGAAAAAGGGGACTGGAGCCAGTTTCAGCTCTGCGAGCAAGAAAGAAAGAAATGAAAAAT 1610
Qy	1929	GCCTGCAAAAGGAGCTCTTAAACCTGTACTAGAAAGTTCCCTCCGAGACAACAGGA 1988
Db	1611	GCCTGCAAAAGGAGCTCTTAAACCTGTACTAGAAAGTTCCCTCCGAGACAACAGGA 1670
Qy	1989	TGCAAGAGAGGACAGATCAAAATTTCCATCATGCAACCCCGGAGCTCACGTGTGGCCGAC 2048
Db	1671	TGCAAGAGAGGACAGATCAAAATTTCCATCATGCAACCCCGGAGCTCACGTGTGGCCGAC 1730
Qy	2049	ACAGGGCCCAAACTGCGAGCTCCGAAATGCACTGGGCTTGGTGTATTTCCCAAGGAAGGC 2108
Db	1731	ACAGGGCCCAAACTGCGAGCTCCGAAATGCACTGGGCTTGGTGTATTTCCCAAGGAAGGC 1790
Qy	2109	TGCAAGATTCGATGTGCTCAACGAGACCGAGCACTGGGAGGAAGCAAGGTGCTCATCTTT 2168
Db	1791	TGCAAGATTCGATGTGCTCAACGAGACCGAGCACTGGGAGGAAGCAAGGTGCTCATCTTT 1850
Qy	2169	GATGACTCTTTGAGCAGCAGGATATGCGAGGATGCCTCATCTTTCCGCTGATATTCATC 2228
Db	1851	GATGACTCTTTGAGCAGCAGGATATGCGAGGATGCCTCATCTTTCCGCTGATATTCATC 1910
Qy	2229	GTGAGTGTGCGATCCGAACTGCACACAGCAGAGCGAGCCTTCCAGCAATTTAG 2288
Db	1911	GTGAGTGTGCGATCCGAACTGCACACAGCAGAGCGAGCCTTCCAGCAATTTAG 1970
RESULT 9		
AF289489		2680 bp mRNA linear PRI 18-DEC-2000
LOCUS		
DEFINITION		Homo sapiens aspartyl beta-hydroxylase 2.8 kb transcript mRNA,
ACCESSION		complete cds; alternatively spliced.
VERSION		AF289489
KEYWORDS		AF289489.1 GI:11878115
SOURCE		Homo sapiens (human)
ORGANISM		Homo sapiens
REFERENCE		1 (bases 1 to 2680)
AUTHORS		Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P., Link, J., O'Neill, K.T., Focht, R.J., Scully, M.S., Hollis, J.M., Hollis, G.F. and Friedman, P.A.
TITLE		Aspartyl beta -hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junctin
JOURNAL		J. Biol. Chem. 275 (50), 39543-39554 (2000)
MEDLINE		20564328
PUBMED		10956665
AUTHORS		2 (bases 1 to 2680)
TITLE		Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and Friedman, P.A.
JOURNAL		Direct Submission
FEATURES		Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA
source		Location/Qualifiers
		1. .2680
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/chromosome="8"
		/map="8q12 at STS WT-11767"
		/tissue_type="liver"
		1. .11
		12. .953
		/note="Humburg; asparaginyl beta-hydroxylase; Asph; non-catalytic isoform; alternatively spliced"
5'UTR		
CDS		



```
source
1. .1080
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3887962"
/tissue_type="Lung, carcinoma, large cell
undifferentiated."
/clone_lib="NIH_MGC_69"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .>1080
/gene="ASPH"
/note="synonyms: JCTN, HAAH, CASQ2BP1, BAH"
/db_xref="LocusID:444"
/db_xref="WIM:600582"
189. .>1080
/gene="ASPH"
/codon_start=1
/product="ASPH protein"
/protein_id="AAH15518.1"
/db_xref="GI:34189305"
/db_xref="LocusID:444"
/db_xref="MIM:600582"
/translation="MAQRKNAKSGNSSSGSGTSAGSSSPARRETKHGHRNG
RKGLSGTSFTFMVITALLGVMTSVAVWFDLVYEEVLKGLIYDADGDGDFDND
AKVLGLKERSSTPEVPPEAEHPTEPEQVPEAEQNIIDEAKEQIQLLHEMYH
AEHVEGDLQOEDGPTGEPQOEDDEFLMATVDVDRFTLRLPEVSHETESHYVEVD
SSEPVVEDRLHDTDDTVYVYEEQVPLENEGIEITEVTAPEDNPVEDSQVIV
EVSIFPVPEQOEVPPEKKK"

gene
CDS
1 CGGACCGTGCATGCGCCAGCGTGAAGATCCAAAGAGCAGCGGCAAGCAGCAGCAGCGG 60
178 CGGACCGTGCATGCGCCAGCGTGAAGATCCAAAGAGCAGCGGCAAGCAGCAGCAGCGG 237
61 GCTCCGCGAGCGGTAGCAGCAGTGGCGCAGCAGCAGCGCCCGCGGCGGAGAGACAA 120
238 GCTCCGCGAGCGGTAGCAGCAGTGGCGCAGCAGCAGCGCCCGCGGCGGAGAGACAA 297
121 AGCATGAGGACACAAAGATGGAGAAAGCGGACTCTCGGGAACTTCATTCTTCAGT 180
298 AGCATGAGGACACAAAGATGGAGAAAGCGGACTCTCGGGAACTTCATTCTTCAGT 357
181 GGTATTGATGATTGCAATGCTGGCGGCTGCGACATCTGTAGCTGCTGTTGGTTGATC 240
358 GGTATTGATGATTGCAATGCTGGCGGCTGCGACATCTGTAGCTGCTGTTGGTTGATC 417
241 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTAGTCTGATGCTGATGGAG 300
418 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTAGTCTGATGCTGATGGAG 477
301 ATTTTGATGATGATGCCAAAGTTTATTAGACTTAAAGAGAGATCTACTTCAGAGC 360
478 ATTTTGATGATGATGCCAAAGTTTATTAGACTTAAAGAGAGATCTACTTCAGAGC 537
361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGCAGAGCAGGTTCTGTGG 420
538 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGCAGAGCAGGTTCTGTGG 597
421 AGGCAGAACCCAGAAATCCGAAGATGAAGCAAGAAAGCAAAATTCAGTCCCTTCCCATG 480
598 AGGCAGAACCCAGAAATCCGAAGATGAAGCAAGAAAGCAAAATTCAGTCCCTTCCCATG 657
481 AAATGGTACAGCGAGAAATGTTGAGGGAGAGAGTTCGCAACAAAGAGATGGACCCACAG 540
658 AAATGGTACAGCGAGAAATGTTGAGGGAGAGAGTTCGCAACAAAGAGATGGACCCACAG 717
541 GAGAACCAACAGAGAGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATG 600

ORIGIN
Query Match 35.7%; Score 829.6; DB 9; Length 1080;
Best Local Similarity 93.6%; Pred. No. 4.2e-169;
Matches 899; Conservative 0; Mismatches 4; Indels 57; Gaps 1;

Qy 1 CGGACCGTGCATGCGCCAGCGTGAAGATCCAAAGAGCAGCGGCAAGCAGCAGCAGCGG 60
Db 178 CGGACCGTGCATGCGCCAGCGTGAAGATCCAAAGAGCAGCGGCAAGCAGCAGCAGCGG 237
Qy 61 GCTCCGCGAGCGGTAGCAGCAGTGGCGCAGCAGCAGCGCCCGCGGCGGAGAGACAA 120
Db 238 GCTCCGCGAGCGGTAGCAGCAGTGGCGCAGCAGCAGCGCCCGCGGCGGAGAGACAA 297
Qy 121 AGCATGAGGACACAAAGATGGAGAAAGCGGACTCTCGGGAACTTCATTCTTCAGT 180
Db 298 AGCATGAGGACACAAAGATGGAGAAAGCGGACTCTCGGGAACTTCATTCTTCAGT 357
Qy 181 GGTATTGATGATTGCAATGCTGGCGGCTGCGACATCTGTAGCTGCTGTTGGTTGATC 240
Db 358 GGTATTGATGATTGCAATGCTGGCGGCTGCGACATCTGTAGCTGCTGTTGGTTGATC 417
Qy 241 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTAGTCTGATGCTGATGGAG 300
Db 418 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTAGTCTGATGCTGATGGAG 477
Qy 301 ATTTTGATGATGATGCCAAAGTTTATTAGACTTAAAGAGAGATCTACTTCAGAGC 360
Db 478 ATTTTGATGATGATGCCAAAGTTTATTAGACTTAAAGAGAGATCTACTTCAGAGC 537
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGCAGAGCAGGTTCTGTGG 420
Db 538 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGCAGAGCAGGTTCTGTGG 597
Qy 421 AGGCAGAACCCAGAAATCCGAAGATGAAGCAAGAAAGCAAAATTCAGTCCCTTCCCATG 480
Db 598 AGGCAGAACCCAGAAATCCGAAGATGAAGCAAGAAAGCAAAATTCAGTCCCTTCCCATG 657
Qy 481 AAATGGTACAGCGAGAAATGTTGAGGGAGAGAGTTCGCAACAAAGAGATGGACCCACAG 540
Db 658 AAATGGTACAGCGAGAAATGTTGAGGGAGAGAGTTCGCAACAAAGAGATGGACCCACAG 717
Qy 541 GAGAACCAACAGAGAGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATG 600
```





```
Qy 901 AAGTAAGCATTTTCTCTGTGAAGACACGACGAGNAGTACCACCGAACAAATAG 955
Db 785 AAGTAAGCATTTTCTCTGTGAAGACACGACGAGNAGTACCACCGAACAAATAG 839

RESULT 15
AR532049
LOCUS AR532049 2442 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6727080..
ACCESSION AR532049
VERSION AR532049.1 GI:53920638
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2442)
AUTHORS Radosevich J
TITLE Gene encoding labyrinthin, a marker for cancer
JOURNAL Patent: US 6727080-A 1 27-APR-2004;
FEATURES
source
1..2442
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 35.2%; Score 819; DB 6; Length 2442;
Best Local Similarity 98.8%; Pred. No. 8.4e-167;
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 121 AGCATGGAGGACACAAAGTGGGAGGAGGCGGACTCTCGGAACTTCATCTTCACGT 180
Db 5 AGCTTGAAGACACAGATGGGAGGAGGCGGACTCTCAGGAACCTTCATCTTCACGT 64

Qy 181 GGTATTATGGTATTCGATTCCTGGGCGTCTGGACATCTGTAGCTGTGTTGGTTGATC 240
Db 65 GGTATTATGGTATTCGATTCCTGGGCGTCTGGACATCTGTAGCTGTGTTGGTTGATC 124

Qy 241 TTGTTGACTATGAGGAAGTCTTAGGAAACTAGGAATCTATGATCTGTAGTGTGATGGAG 300
Db 125 TTGTTGACTATGAGGAAGTCTTAGGAAACTAGGAATCTATGATCTGTAGTGTGATGGAG 184

Qy 301 ATTTTGATGTGGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC 360
Db 185 ATTTTGATGTGGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC 244

Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 420
Db 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 304

Qy 421 AGGCAGAACCCCGAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480
Db 305 AGGCAGAACCCCGAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 364

Qy 481 AAATGGTACACGACAAACATGTTGAGGGAGAGACTTGCACCAAGAGATGGACCCACAG 540
Db 365 AAATGGTACACGACAAACATGTTGAGGGAGAGACTTGCACCAAGAGATGGACCCACAG 424

Qy 541 GAGAACCAACACAGAGGATGATGATTTCTTATGGCGACTGTAGTGTAGATGATGTTG 600
Db 425 GAGAACCAACACAGAGGATGATGATTTCTTATGGCGACTGTAGTGTAGATGATGTTG 484

Qy 601 AGACCTTGGAACTGAAATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAAGAGA 660
Db 485 AGACCTTGGAACTGAAATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAAGAGA 544

Qy 661 CAGTTTCAAGAAGCTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720
Db 545 CAGTTTCAAGAAGCTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 604

Qy 721 ATTCAGTGAACAGTAGTAGAAGATGAAGATTGCCACCATGATACAGATGATGTAACAT 780
Db 605 ATTCAGTGAACAGTAGTAGAAGATGAAGATTGCCACCATGATACAGATGATGTAACAT 664
```

```
Qy 781 ACCAAGTCTATGAGGAAACAAGCAGATATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840
Db 665 ACCAAGTCTATGAGGAAACAAGCAGATATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 724

Qy 841 CAGAAGTAACCTGCTCCCTGAGGATAATCCTGTAGAAGATTACAGGTAATTTGTAGAAG 900
Db 725 CAGAAGTAACCTGCTCCCTGAGGATAATCCTGTAGAAGATTACAGGTAATTTGTAGAAG 784

Qy 901 AAGTAAGCATTTTCTCTGTGGAAGACACGACGAGAGTACCACCGAACAAATAG 955
Db 785 AAGTAAGCATTTTCTCTGTGGAAGACACGACGAGAGTACCACCGAACAAATAG 839
```

Search completed: March 25, 2005, 12:32:14  
Job time : 9726 secs

10-15-2010 2:20 PM





Db 1 ATGCCCCAGCGTAAAGTCCAAAGAGCAGCGGCAACAGCAGCAGCAGCGGCTCCGCGAGC 60  
Qy 72 GGTAGCACGAGTGGCGCAGCAGCAGCGCCCGGGGCCCGGAGAGAGACAAAGCATGAGGA 131  
Db 61 GGTAGCACGAGTGGCGCAGCAGCAGCGCCCGGGGCCCGGAGAGAGACAAAGCATGAGGA 120  
Qy 132 CACAAGAAATGGGGAAGGCGACTCTCGGGAACCTTCACTTTCACGTGTTTATGGTG 191  
Db 121 CACAAGAAATGGGGAAGGCGGACTCTCAGGAACCTTCACTTTCACGTGTTTATGGTG 180  
Qy 192 ATTGCATTTGTCGGCGCTCGACACTCTGTAGCTGTGTTTGGTTGATCTTTGTGACTAT 251  
Db 181 ATTGCATTTGTCGGCGCTCGACACTCTGTAGCTGTGTTTGGTTGATCTTTGTGACTAT 240  
Qy 252 GAGGAAGTTCTAGGAAAATAGGAATCTATGATGCTGATGGTGTGAGAGATTTTGTGTG 311  
Db 241 GAGGAAGTTCTAGNN 300  
Qy 312 GATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCG 371  
Db 301 NNN 360  
Qy 372 CCAGAAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGGAGGCGAAGACC 431  
Db 361 CCAGAAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGGAGGCGAAGACC 420  
Qy 432 CAGAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAAATGGTACAC 491  
Db 421 CAGAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAAATGGTACAC 480  
Qy 492 GCAGAACATGTTGAGGGAAGAGACTTGCAACAAAGATGAGCCACAGGAGAAACCAAA 551  
Db 481 GCAGAACATGTTGAGGGAAGAGACTTGCAACAAAGATGAGCCACAGGAGAAACCAAA 540  
Qy 552 CAAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATGATGATGATG 611  
Db 541 CAAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATGATGATGATG 600  
Qy 612 CTTGAAGTATCTCATGAAGAACCGAGCATAGTTTACACGTTGGAAGAGACAGTTTCAAA 671  
Db 601 CTTGAAGTATCTCATGAAGAACCGAGCATAGTTTACACGTTGGAAGAGACAGTTTCAAA 660  
Qy 672 GACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAGATTCAGTGAA 731  
Db 661 GACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAGATTCAGTGAA 720  
Qy 732 CCAGTAGTAGAAGATGAAGATTGCACCATGATGATGATGATGATGATGATGATGATGAT 791  
Db 721 CCAGTAGTAGAAGATGAAGATTGCACCATGATGATGATGATGATGATGATGATGATGAT 780  
Qy 792 GAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAAATCACAAGAGTAACT 851  
Db 781 GAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAAATCACAAGAGTAACT 840  
Qy 852 GCTCCCTCTGAGGATAAATCTGTAGAGATTCAAGGTAATTTGTAGAGAAAGTAAAGCAT 911  
Db 841 GCTCCCTCTGAGGATAAATCTGTAGAGATTCAAGGTAATTTGTAGAGAAAGTAAAGCAT 900  
Qy 912 TTTCTGTGGAAGAACAGCAGAGAGTACCAACAGAAAACAAATAGAAAACAGATGATCCA 971  
Db 901 TTTCTGTGGAAGAACAGCAGAGAGTACCAACAGAAAACAAATAGAAAACAGATGATCCA 960  
Qy 972 GAACAAAAGCAAAAGTTTAAAGAAAAGAGCCCTTAAATTTTAAATTTTAAAGACT 1031  
Db 961 GAACAAAAGCAAAAGTTTAAAGAAAAGAGCCCTTAAATTTTAAATTTTAAAGACT 1020  
Qy 1032 ATTTAAAGCTGAATCTGATGCTCGAGAAAACCTCCGTAAAGGGGAAAATTTGAGGAAGCA 1091  
Db 1021 ATTTAAAGCTGAATCTGATGCTCGAGAAAACCTCCGTAAAGGGGAAAATTTGAGGAAGCA 1080  
Qy 1092 GTGAATGCAATTTAAAGAACTAGTACCAATACCTTCAGATCCAGCAGCAAGATATGGG 1151  
Db 1081 GTGAATGCAATTTAAAGAACTAGTACCAATACCTTCAGATCCAGCAGCAAGATATGGG 1140

Qy 1152 AAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGTAATGAGGTGCTACGTGA 1211  
Db 1141 AAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGTAATGAGGTGCTACGTGA 1200  
Qy 1212 GCCATCGAGACTTCAAGAGGTGGCCAGCTTACCTGATGTCCTCCGACAGCTGCTGAG 1271  
Db 1201 GCCATCGAGACTTCAAGAGGTGGCCAGCTTACCTGATGTCCTCCGACAGCTGCTGAG 1260  
Qy 1272 CTGAGTTTGAAGCGTCCGCTCAGACAGGCAACAAATTTCTAGGTATATGAGAGGTTCCCTG 1331  
Db 1261 CTGAGTTTGAAGCGTCCGCTCAGACAGGCAACAAATTTCTAGGTATATGAGAGGTTCCCTG 1320  
Qy 1332 CTTTACCTCGAGAGATTTAGTTTCAACTTATTTCCCAATGATCTTCCCTTAAAAAATGACCTT 1391  
Db 1321 CTTTACCTCGAGAGATTTAGTTTCAACTTATTTCCCAATGATCTTCCCTTAAAAAATGACCTT 1380  
Qy 1392 GGGCTGGGATACCTCTTGTATAGAGATATGCAATGCAAGAAAGATTTATGAAGAGTG 1451  
Db 1381 GGGCTGGGATACCTCTTGTATAGAGATATGCAATGCAAGAAAGATTTATGAAGAGTG 1440  
Qy 1452 CTGAGTGTGACACCTTAATGATGCTTAAAGTCCATTTATGGCTTCACTCCTGAAGGCA 1511  
Db 1441 CTGAGTGTGACACCTTAATGATGCTTAAAGTCCATTTATGGCTTCACTCCTGAAGGCA 1500  
Qy 1512 CAGAACAAAATTTGCTGAGAGCATCCCATATTTAAAGGAAAGGAATAGAATCCCGAGATCCT 1571  
Db 1501 CAGAACAAAATTTGCTGAGAGCATCCCATATTTAAAGGAAAGGAATAGAATCCCGAGATCCT 1560  
Qy 1572 GGCATGTATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGAGGTTGGAAAC 1631  
Db 1561 GGCATGTATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGAGGTTGGAAAC 1620  
Qy 1632 AAAGAGCATAATAAGTGGTATGAGCTTGGGCAACAAGAGAGGACACTTTGCTGCTGG 1691  
Db 1621 AAAGAGCATAATAAGTGGTATGAGCTTGGGCAACAAGAGAGGACACTTTGCTGCTGG 1680  
Qy 1692 CAACGCTCACTCTACAATGGAATGGAATGGAAGCAACGCTTGGTGGACCCCAAGAA 1751  
Db 1681 CAACGCTCACTCTACAATGGAATGGAATGGAAGCAACGCTTGGTGGACCCCAAGAA 1740  
Qy 1752 ACCGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAA 1811  
Db 1741 ACCGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAA 1800  
Qy 1812 GGCCTTGCAGTGTGATAAAGCAAGGTCCTCTCTGCTGAGGATGAAAACCTGAGG 1871  
Db 1801 GGCCTTGCAGTGTGATAAAGCAAGGTCCTCTCTGCTGAGGATGAAAACCTGAGG 1860  
Qy 1872 GAAAAGGGGACTGGAGCCAGTTTCAACGCTGTGGCAGCAAGAAAGATGAAAATGCC 1931  
Db 1861 GAAAAGGGGACTGGAGCCAGTTTCAACGCTGTGGCAGCAAGAAAGATGAAAATGCC 1920  
Qy 1932 TGCAAAAGGAGCTCTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGACAACAGGATGC 1991  
Db 1921 TGCAAAAGGAGCTCTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGACAACAGGATGC 1980  
Qy 1992 AGAAGAGACAGATCAAAATTTCCATCATGCAACCCCGGAGCTCACTGTGTGGCGGCACACA 2051  
Db 1981 AGAAGAGACAGATCAAAATTTCCATCATGCAACCCCGGAGCTCACTGTGTGGCGGCACACA 2040  
Qy 2052 GGGCCCAAAACTGCGAGCTCCGATGCTGAGTGGCTTGGTGTATCCCAAGGAGGCTGC 2111  
Db 2041 GGGCCCAAAACTGCGAGCTCCGATGCTGAGTGGCTTGGTGTATCCCAAGGAGGCTGC 2100  
Qy 2112 AAGATTGATGTGCAACAGGAGCA 2136  
Db 2101 AAGATTGATGTGCAACAGGAGCA 2125







QY	2040	TGGCCGCACACAGGGCCCAACAACTGCAGGCTCGAATGCACCTGGGCTTGGTGAATCCC	2099
Db	2154	TGGCCGCATACAGGACCCCAAACTGCAGGCTCGAATGCATCTGGGGTTAGTGATCCCC	2213
QY	2100	AAGGAAGGCTGCAAGATTTCGATGTCACACGAGACAGGACCTGGGAGGAAGCAAGGTG	2159
Db	2214	AAGGAAGGCTGCAAGATCCGGTGTGCCAATGAGACCAAGGACGTGGGAAGCAAGGTG	2273
QY	2160	CTCATCTTTGATGACTCTTTTGAGCAGAGGTATGGCAGGATGCTCATCTTTCCGGCTG	2219
Db	2274	CTCATCTTTGATGACTCTTTTGAGCAGAGGTATGGCAGGATGCTCTTTCCGGCTG	2333
QY	2220	ATATTTCATCTGGATGTGTGGCATCCGGAATCGACACACAGCAGACGACGCTTTCCA	2279
Db	2334	ATATTTCATCTGGATGTGTGGCATCCGGAATCGACACACAGCAGACGACGCTTTCCA	2393
QY	2280	GCAATTAGCATGAATTCATGCAAGCTTGG	2309
Db	2394	GCAATTGAAAGGCACCTGACGCACTTGG	2423
RESULT 4			
LOCUS	AK049506	3924 bp mRNA linear	HTC 03-APR-2004
DEFINITION	Mus musculus 7 days embryo whole body cDNA, RIKEN full-length enriched library, clone:Q430018B11		
	product:aspartate-beta-hydroxylase, full insert sequence.		
ACCESSION	AK049506		
VERSION	AK049506.1	GI:26340239	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;		
REFERENCE	1		
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,Y., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE	5		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		

6 (bases 1 to 3924)  
REFERENCE  
AUTHORS  
Adachi,J., Aizawa,K., Akimura,T., Arikawa,T., Bono,H., Carninci,P.,  
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,T., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
Okazaki,Y., Saito,R., Shihata,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.jp,  
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.jp/  
URL:http://fantom.gsc.riken.jp/  
FEATURES  
source  
Location/Qualifiers  
1..3924  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:C430018B11"  
/db\_xref="taxon:10090"  
/clone="C430018B11"  
/tissue type="whole body"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="7 days embryo"  
177..2354  
/notes="unnamed protein product; aspartate-beta-hydroxylase  
(MGD)MG1:1914186, GH|NM\_023086, evidence: BLASTN, 99%,  
match=4404)  
putative"  
/codon\_start=1  
/protein\_id="BAC33783.1"  
/db\_xref="GI:26340240"  
/translation="MAPRKNAKGGGSSSSSGSGSGSPSTGSGSSSSSPGARRE  
KHGDFGKRRGIGSGSFTFWFMIALIGVMTSVAVFDLYDEVILGKVEYKAD  
QGDGFDVDDKVLGKSPPTPEEAEATHAELEEAPQEGADIQNVEDVEVQD  
IOSLQESVHTDHLADGLAGBPQEVDFLTVDSDRFEDELPQTHVEEIDTVH  
VEDTASQHPNDMEETNPEVSEVHODYDEPVPSEHGEVAISDNTIDSSIIIS  
EETINVASVEEQDTTPVKKKKPKLLNKDKTIKABLDAAEKLKRKGKIEAVNAPEEL  
VRKPSQPRARYGAQCEDDLAEKQSNVEIRLAETIYQEAADLPAPDTDLVKLSLR  
RSRQOFLGMRGSLITLQRLVQFPSTDITLKNLDGVGHLLGDNDSAKKVEVSLNV  
TPNDGKAVHYGFIKQAQNIISPIYKLEGISGPGTDDGRGFYHLGDMQVRGNK  
EAYKVELGHGRGHFAVSQRSILYVNLGKQAQPPWTRETGYTELVKSLERNWKLIR  
EGLMWMDKAGLFLPDENLREKQDSQPTLWOQGRKNENACKGAPKTCALLKFESET  
TCRRGQIKYISIMHPGTIVWPHGTIGRNLMLHGLVLPKEGCKIRCANCETRTWEECK  
VLIFDDSFHEVWQDASSFRLIFIVDWHPELTPQORRSLPAI"

ORIGIN  
Query Match 59.1%; Score 1374.4; DB 3; Length 3924;  
Best Local Similarity 78.1%; Pred. No. 0;  
Matches 1805; Conservative 0; Mismatches 361; Indels 144; Gaps 7;  
Qy 15 GCCACGCGTAAAGTGCACAGACGCGGCAACAGCAGCAGCGCGCTCGGCAGCGGT 74  
Dd 195 GCTAAGGGCGCGCGGCACACAGCAGCAGCGGCGCAGCGGCTCGGCAGCGGT 254  
Qy 75 AGCACGAGTGCGGCGCAGCAGC-----AGCCCCGGGCCCCGAGAGACACAAG 122

Db 255 AGCCGAGCAGCGGTAGCAGCGGTAGCAGCGGTAGCAGCGCGGCGCTCGGAGAGAGGCAAG 314  
Qy 123 CATGGAGGACACAAGAAATGGAGGAAAGCGGAGCTCTCGGGAACTTCAATCTTCAGGTGG 182  
Db 315 CACGGAGGACACAAGAAATGGAGGAAAGCGGAGGATTCAGGAGGCTCTTTTTCACATGG 374  
Qy 183 TTTATGGTGAATGCATTTGCTGGGCGCTCTGGACATCTGTAGCTGTCTGTGTTGGTTGATCTT 242  
Db 375 TTTATGGTGAATGCATTTGCTGGGCGCTCTGGACATCTGTGCTGTCTGTGTTGGTTGATCTT 434  
Qy 243 GTTCACTATCAGGAAGTTCTAGAAAACTAGGAATCTATGATGCTGATGCTGATGATGAGAT 302  
Db 435 GTTCATTTAGGAAGTTCTAGGAATCTAGGAGTCTATGATGCGGATGGCATGGAGAC 494  
Qy 303 TTTGATGTGATGATGCCAAAGTTTATTTAGGACTTTAAGAGAGATCTACTTCAGAGCA 362  
Db 495 TTTGATGTGATGATGCCAAAGTTTATTTAGGCTTTAAGAAAGATCTCTCTCTGAGAGG 554  
Qy 363 GCATGCTCCGCC---AGAGAGGCTGAGCCACACACTGAGCCGAGGAGAGGTTCCTGTG 419  
Db 555 ACATTTCCACCGGAGGAGGAGGAGAGACTCACGCTGAGCTGGAGGAGCAGGCCCTCGAG 614  
Qy 420 GAGGAGAACCCGAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCAT 479  
Db 615 GCGGCAGACATCCAGAAATGTTGAAGATGAAGTAAAGGAAACAAATTCAGTCCCTTCTTCAG 674  
Qy 480 GAAATGTTACACGACGACATGTTGAGGGAGAGACTTGCAACAGAGAGATGAGCCACCA 539  
Db 675 GAATCAGTACACAGACCATG-----ACCTTGAGACAGATGGGTGGCA 719  
Qy 540 GGAGAACCAACAAGAGATGATGATGTTCTTATGGGACTGATGTAGATGATGATTT 599  
Db 720 GGAGAACCAACAGCGGAGGTTGAGGACTTCTCTACAGTGACCCGACAGTATGACAGATTT 779  
Qy 600 GAGACCTGGAACCTGAAATCTCTCATGAAGAAACCGAGCATAGTTTACCACGCTGGAAGAG 659  
Db 780 GAGGACCTGGAACCCGGGACAGTTTCATGAAGAAATTCAGGAGTACTTACCATGTGGAAGAC 839  
Qy 660 ACAGTTTTCACAGACTGTAATCAGGATATGAAGAGATGATGCTGAGCAGAGAAATCCA 719  
Db 840 ACAGCATCGCAGAACCATCCAATGACATGAGAGAGATGACGAATGAGCAGGAATTT-- 897  
Qy 720 GATTCAGTGAACCATGATAGAAATGAAGATGACCATGATACAGATGATTAACA 779  
Db 898 -----CAGAGAGTAAAG 911  
Qy 780 TACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATC 839  
Db 912 CACCAAGACTATGATGAAC---CAGTATATGAACCCCTCAGAGCATGAAGGGGTGCGGAT 968  
Qy 840 ACAGAGTAATCTGCTCCCTCAGGATTAATCTGTAGAAGATTCACAGGTAAATTTAGAA 899  
Db 969 TCA-----GATAACACCATAGATGATTCAGCATATATCTCAGAA 1007  
Qy 900 GAAGTAAGCATTTTCTCTGGAAGAACAGCAGAGAGTACCACACAGAAACAAATAGAAA 959  
Db 1008 GAATATAATCTGCTCTCTGTTGAAGAACAGCAGCAGACACACACC----- 1051  
Qy 960 ACAGATGATCCAGAACAAAAGTTAAGAAAAGAGGCTTAACCTTTTAATATAA 1019  
Db 1052 -----AGTTAAGAAAGAAAGGCTTAACCTCTGTAACAAA 1085  
Qy 1020 TTTGATTAAGACTATTAAGCTGAACCTTGATCTGACAGAAAACTCCGTTAAAAAGGGAAAA 1079  
Db 1086 TTTGATTAAGCAATTAAGGCTGAGCTGGATGCTGCAGAAAAGCTCCGAAAAAGGGGTAAA 1145  
Qy 1080 ATTGAGGAACAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCAGCA 1139  
Db 1146 ATTGAGGAACAGTGAATGCAATTTAAAGAACTAGTTGGAAGTACCTCAGAGCCACGA 1205  
Qy 1140 GCAGATATGGGAGGCGCAGTCTGAGGATGATTTGGCTGAGAGGAGGAGATGATGAG 1199  
Db 1206 GCAAGATATGGCAAGCGGAGTGTGAAGATGATCTTGGCAGAGAGCAGAGAGCAATGAG 1265

Qy 1200 GTGCTACGTGGAGCCATCGAGACCTTACCAAGAGGTGGCGAGCCTTACTGTATGCTCCTGCA 1259  
Db 1266 GTTCTGGCGAGGGCCATCGAGACCTTACCAAGAGGTGGCGAGCCTTACTGTATGCTCCTGCA 1325  
Qy 1260 GACCTGCTGAAGCTGAGTTTGAAGGCTCGCTCAGACAGGCAAAATTTCTAGTCTCATATG 1319  
Db 1326 GACCTGCTGAAGCTGAGTTTGAAGGCTCGGAAACGGCAGCAGTTTCTAGTCTCATATG 1385  
Qy 1320 AGAGTTTCCCTGCTTACCTCGCAGAGATTTAGTTTCAACTATTTTCCCAATGATCTCTCTTA 1379  
Db 1386 AGAGTTTCTCTACTTACCTTCAGAGACTAGTTTCACTGTCTTCCCTAGTACTACTCTTA 1445  
Qy 1380 AAAAATGACCTTGGCGTGGATACCTCTGTATAGAGATATGACATGCAAAAGATTT 1439  
Db 1446 AAAAACGACCTTGGCGTAGATACCACTTGTGGAGACAAACGACGTGCCAAGAAGTT 1505  
Qy 1440 TATGAAGAGTGTGATGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGCTTTC 1499  
Db 1506 TACGAAGAGTGTAAATGTGACACCAATGATGGCTTCGCTAAAGTGCATTAACGGCTTC 1565  
Qy 1500 ATCTGTAAGGCACAGAACAAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAA 1559  
Db 1566 ATCTGTAAGGCACAGAACAAAGATATCTGAGAGCATTCCTACTTTAAAGGAAGGAATCGAA 1625  
Qy 1560 TCCGGAGATCTCGCACTCATGATGGAGATTTTATTTCCACCTGGGGATGCCATGCAG 1619  
Db 1626 TCTGGGACCTTCGCAAGGATGATGGCGGTTTACTTCCACTTGGGGATGCCATGCAG 1685  
Qy 1620 AGGTTGGGAAACAAAGAGCATAATAGTGGTATGAGCTTGGGCACAAAGAGAGACACTTT 1679  
Db 1686 AGGTTGGGAAACAAAGAGCATAATAGTGGTATGAACTTGGGCACAAAGAGAGACACTTT 1745  
Qy 1680 GCATCTGCTGGCAACCGCTCACTCTACAATGTGAATGAGTGAAGTCTTTAGAAAGAACTTGAAGTTA 1739  
Db 1746 GCCTCTGCTGGCAGCGCTTCCCTCTACAATGTGAATGGTCTGGAAGGCTCAGCGCTGGTG 1805  
Qy 1740 ACCCAAAAGAAACCGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTA 1799  
Db 1806 ACACCCAGGAGAGCTGGCTACACAGAGCTAGTGAAGTCTTTAGAGAGAACTTGAAGTTA 1865  
Qy 1800 ATCCGAGTGAAGCGCTTCCAGTGTGATGATGAAGCCAAAGGTCTCTCTGCTGCTGAGAT 1859  
Db 1866 ATCCGATGAAGCGCTCTAGTGTGATGATGAAGCCAAAGGTCTCTCTGCTGCTGAGHAC 1925  
Qy 1860 GAAAACTGAGGAAAAAAGGAGCTGGAGCCAGTTTCAAGTGTGGCAGCAAGGAAGAAGA 1919  
Db 1926 GAAAACTTCCGGAGAAAGGGCGACTGGAGCCAGTTTCACTGTGGCAGCAAGGAAGAAG 1985  
Qy 1920 AATGAAATGCTTCGAAAGGAGCTCTTAAACCTGTACTTCTTACTAGAAAAGTTTCCCGAG 1979  
Db 1986 AATGAGAAATGCTTGAAGGAGCGCTTAAAGCTGTGCTTACTAGAAAAGTTTCCGAA 2045  
Qy 1980 ACAACAGATGAGAGAGAGCAGATCAATATTTTCCATCATGCAACCCCGGACTCACGTG 2039  
Db 2046 ACAACAGATGAGAGAGAGCAGATCAATATTTTCCATCATGCAACCTTGAAGTCTCATGTG 2105  
Qy 2040 TGGCCGACACAGGCGCCCAAACTGAGGCTCCGAATGCACTGGGCTTGGTGTATCC 2099  
Db 2106 TGGCCGATACAGGACCCCAAACTGAGGCTCCGAATGCACTTGGGTTAGTGTATCC 2165  
Qy 2100 AAGAGGCTGCAAGATTCGATGTCCCAACGAGACCAAGGACTTGGGAGGAAGCAAGGTG 2159  
Db 2166 AAGAGGCTGCAAGATTCGATGTCCCAATGAGACCAAGGACTTGGGAGGAAGCAAGGTG 2225  
Qy 2160 CTCACTTTTGTGATGACTCTTTTTCAGCAGCAGGTATGAGCAGGATGCTCATCTTTTCCGGCTG 2219  
Db 2226 CTCACTTTTGTGATGACTCTTTTTCAGCAGCAGGTATGAGCAGGATGCTCATCTTTTCCGGCTG 2285  
Qy 2220 ATATTCACTGTGATGCTGTGGCATCCGAACTGTACACCAAGCAGCAGGAGCGACCTTCCA 2279  
Db 2286 ATATTCACTGTGATGCTGTGGCATCCGAACTGTGGCAGGCTGAGCCCTCAGCAGAGAGCGACCTTCCC 2345





QY 471 CTTCTCCATGAATGGTACACGACACATGTTGAGGAGAGACTTTCACACAGAGAT 530  
Db 551 CTTCTTCAGGATCACTACACAGACCATG-----ACCTTGAGCAGAT 595  
QY 531 GGACCCACAGAGAACCAACACAGAGAGATGATGATTTCTTATGCGCAGCTGATGATAG 590  
Db 596 GGGCTGGCAGAGAACCAACAGCGAGGTTGAGGACTTCTTCACAGTGCACCGCAGATGAT 655  
QY 591 GATGATTTGAGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACAC 650  
Db 656 GACAGATTTGAGGACCTGGAAACCCGGGACAGTTTCATGAAGAAATCGAGGATACTTACCAT 715  
QY 651 GTGAAGAGACAGTTTCACAAAGACTGTATCAGGATATGGAAGAGATGATGTCGAGCAG 710  
Db 716 GTGAAGAGACACAGCATCGCAGAACCATCCAAATGACATGGAAGAGATGACGAATGAGCAG 775  
QY 711 GAAATCCAGATTCAGTGAAACCATGATAGTACAGATGAAGATTGACCATGATACAGAT 770  
Db 776 GAAATTCAGAG-----787  
QY 771 GATGTAAACATACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTAGAAATGAAAGG 830  
Db 788 GAAATGAAACCAAGACTATGATGAAC---CAGTATATGAACCCCTCAGAGCATGAAAGG 844  
QY 831 ATGAAATATCAGAAAGTAACTGCTCCCTGAGGATATCTCTGAGAAAGATTACAGGTA 890  
Db 845 GTCGCGATTTCA-----GATAACCATAGATGATTCAGCATA 883  
QY 891 ATTCTAGAAGATGAGCATTTTCTCTGGAAGAACAGCAGAGTATACCAAGAGG 950  
Db 884 ATCTCAGAGAAATTAATGTGCTCTCTGTTGAGAAACAGACAGACACAC 936  
QY 951 AATGAAAAACAGATATCCAGAACAAAAGTTAAGAAAAGAGGCTTAACTT 1010  
Db 937 -----AGTTAAGAAAGAGGCTTAACTT 961  
QY 1011 TTAATTAATTTGATGAAGACTATTAAGCTGAACCTTATGCTCGAGAAACCTCGTAAA 1070  
Db 962 CTGAACAAATTTGATGAAGCAATTAAGCTGAGCTGGATGCTCGAGAAAGTTCGGAAA 1021  
QY 1071 AGGGGAAAAATTCAGGAAGCAGTGAATGCTTTAAAGAACTAGTAGCAAAATACCTCAG 1130  
Db 1022 AGGGTAAATTCAGGAAGCAGTGAACCAATTTGAAGAACTGTTTGAAGTACCTCAG 1081  
QY 1131 AGTCCACGACAGATATGGGAAGGCGCAGTGTGAGATGATTTGCTGAGAGAGGAGA 1190  
Db 1082 AGCCACAGACAGATATGGCAAGCGCAGTGTGAAGTACTTTGCGAGAGAGGAGAGA 1141  
QY 1191 AGTAAAGAGTGCTAGTGGAGCCATCGAGACCTTACCAAGAGGTGGCCAGCCTACCTGAT 1250  
Db 1142 AGCAATGAGGTTCTGCGCAGGCGCCATCGAGACCTTACCAAGAGGCGCCAGCCTGCTGAT 1201  
QY 1251 GTCCCTGCAGACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTA 1310  
Db 1202 GCCCCTACAGACCTGCTGAAGCTGAGCTTGAAGCGAAGTTCGAAACGCGCAGCTTTCTA 1261  
QY 1311 GGTCAATGAGAGTTTCCCTGCTTACCTGCGAGAGATTAAGTCAACTATTTCCCAATGAT 1370  
Db 1262 GGTCAATGAGAGTTTCTTACTTACCTCCAGAGACTAGTTCAACTGTTCCCTAGTAT 1321  
QY 1371 ACTTCTTAATAAATGACCTTGGCTGGGATACCTCTTGTAGAGAGATTAATGACATGCA 1430  
Db 1322 ACTACTTTAAAAACGACCTTGGCTAGGATACCTCTTGTGGAGAACACAGAGTGCC 1381  
QY 1431 AAGAAAGTTTATGAAGAGTGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCAT 1490  
Db 1382 AAGAAAGTTTACGAAGAGTGTCTAATGTGACACCAATGATGGCTTCGCTAAAGTGAT 1441  
QY 1491 TATGGCTTCACTCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCATATTTAAAGAA 1550  
Db 1442 TAGGGCTTCACTCTGAAGGCACAGAAACAAATTCCTGAGAGCATTCCTTACTTAAAGAA 1501

QY 1551 GGAATAGAAATCCGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGAT 1610  
Db 1502 GGAATCGAATCTGGGACCTTGGCAGGATGATGCGGGTTTACTTTCCACTTTGGGGAT 1561  
QY 1611 GCCATGACAGAGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGA 1670  
Db 1562 GCCATGACAGAGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAACTTGGGCAACAAGAGA 1621  
QY 1671 GGACACTTTTCATCTCTGTGGCAAGCGCTCACTCTACAAATGTGAATGAGCTGAAAGACACAG 1730  
Db 1622 GGACATTTGCTCTGTCTGGCAGCGTTCCCTCTACAATGTGAATGCTTGAAGGCTCAG 1681  
QY 1731 CTTTGTGGACCCCAAAAGAAAGCGGCTACACAGAGTTAGTAAAGTCTTTTAGAAGAAAC 1790  
Db 1682 CCGTGTGGACACCCAGGGAGACTGGCTACACAGAGCTAGTGAAGTCTTTTAGAGAGAAAC 1741  
QY 1791 TGGAGTTAATCCGAGATGAAGCGCTTGCAGTGAATGAATGAAGCAAAAGGCTCTTTCCTG 1850  
Db 1742 TGGAGTTAATCCGATGAAGCGCTCATGTGTGAATGAATGAAGCAAAAGGCTCTTTCCTG 1801  
QY 1851 CCTGAGATCAAAACCTGAGGAAAGGAGGAGCTGGAGCCAGTTCAAGCTGTGGCAGCAA 1910  
Db 1802 CCTGAGACGAAACCTTCCGGAGAAAGGCGACTGGAGCCAGTTCACTGTGGCAGCAA 1861  
QY 1911 GGAAGAAAGAAATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACTTACTTAGAAGAA 1970  
Db 1862 GGAAGAAAGAAATGAGATGCTGTAAAGAGGCGCTTAAGACCTGTGCTTTACTAGAAAG 1921  
QY 1971 TTCCCGGAGACACAGATGACAGAGAGACAGATCAATATTTCCATCATGCACCCCGG 2030  
Db 1922 TTTTCCGAAACACAGATGACAGAGAGACAGATCAATATCTCCATCATGACACCTGGA 1981  
QY 2031 ACTCAGCTGTGGCGCACAGAGGCGCAAACTGCAAGCTCCGAATGCACTTGGGCTTG 2090  
Db 1982 ACTCATGTGTGGCGCATACAGGACCCACAACTGCAAGCTCCGAATGCACTTGGGCTTA 2041  
QY 2091 GTGATTTCCAAAGAGGCTGCAAGATTCGATGTGCCAACGACAGACAGGAGGAA 2150  
Db 2042 GTGATTTCCAAAGAGGCTGCAAGATTCGATGTGCCAACGACAGGAGGAA 2101  
QY 2151 GGCAAGTGTCTCATCTTTGATGACTCTTTGAGCAGCAGGTATGGAGATGCTCATCT 2210  
Db 2102 GGCAAGTGTCTCATCTTTGATGACTCTTTGAGCAGCAGGTATGGAGATGCTCATCT 2161  
QY 2211 TTCCGCTGATATTCATGCTGATGTGGCATCCGAACTGACACACAGCAGAGAGC 2270  
Db 2162 TTCCGCTGATATTCATGCTGATGTGGCATCCGAACTGACACACAGCAGAGAGC 2221  
QY 2271 AGCTTCCAGCAATTTAGCATGAATTCATGCAAGCTTGG 2309  
Db 2222 AGCTTCCGCAATTTGAAGGCACTGACGAGACTTGG 2260

## RESULT 6

AY405269  
LOCUS Mus musculus ASPH gene, partial sequence, 2153 bp DNA linear GSS 16-DEC-2003  
DEFINITION Mus musculus ASPH gene, partial sequence, genomic survey sequence.  
ACCESSION AY405269.  
VERSION AY405269.1 GI:39761243  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 2153)  
AUTHORS Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,  
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
Perrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)

```
PUBMED 14671302
REFERENCE 2 (bases 1 to 2153)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2153
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2153
/gene="ASPH"
/locus_tag="HCM2167"
ORIGIN
Query Match 54.4%; Score 1264; DB 9; Length 2153;
Best Local Similarity 69.5%; Pred. No. 5.6e-313; Indels 6; Gaps 2;
Matches 1465; Conservative 0; Mismatches 638;
QY 32 CAAGAGCAGCGGCAACAGCAGCAGCGGCTCCGGCAGCGGTAGCACGAGTGGCGGCGAG 91
DB 48 CAGCGGAGCGGCGAGCGGCTCGGCGAGCGGTAGCCGAGCACGGGTAGCAGCGGTAGCAG 107
QY 92 CAGCAGCCCCGGGCGCGGAGAGAGACAAGCATGAGGAGACACAAGATGGAGGAAGG 151
DB 108 CAGTAGCCCCGGGCTCGAGAGAGGCAAGACACGAGGAGACACAAGATGGAGGAAGG 167
QY 152 CGGACTCTCGGGAACCTCATCTTCACTGCTGCTGTTATGCTGATTCATTCCTGGCGCTG 211
DB 168 AGGGAATTCAGGAGGGTCTTTTTCATGATGTTTATGTCATTCATTCCTGCGGCTG 227
QY 212 GACATCTGAGTGTGCTGTTGTTGATCTTGTGTTGACTATCAGGAAGTTCCTAGGAAACT 271
DB 228 GACATCTGTGCTGTGCTGTTGTTGATCTTGTGTTGATTCATTCATTCAGGAAGTTC 287
QY 272 AGGAATCTATGATGCTGATGATGAGATTTTGAATGATGATGATGATGATGATGATGAT 331
DB 288 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 347
QY 332 AGGACTTAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCC---AGAAGAGGCTGAGCC 388
DB 348 NNGCCTTAAGAAAGATCTCTTCTGAGAGACATTTCCACCGGAGGAGGAGGAGAGAC 407
QY 389 ACACACTGAGCCCGAGGAGCAGGTTCTCTGTGGAGGCAGAACCCAGAAATATCGAAGATGA 448
DB 408 TCACGCTGAGCTGGAGGAGCAGGCGCCCTGAGGGGCGACACATCCAGAAATGTTGAAGATGA 467
QY 449 AGCAAAAGAACAAATTCAGTCCCTTCTCATGAAATGTTACACGAGACATGTTGAGGG 508
DB 468 AGTAAAGGAACAAATTCAGTCCCTTCTTCAGGAATTCAGTACACACAGACCATTCCTGT 527
QY 509 AGAAGACTTGCAACAAGAGATGGACCCACAGGAGAACCAACAAGAGGATGATGAGTT 568
DB 528 TGAAGAC---CTTGAAGCAGATGGCTGGCAGGAGAACCAACGCGGAGGTTGAGGACTT 584
QY 569 TCTTATGGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
DB 585 CCTCAGTGACCGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 644
QY 629 AGAAACCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
DB 645 AGAAATTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 704
QY 689 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
DB 705 GGAAGAGATGACGAATGACGAGGAAATTCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 764
```

Db 1845 TAAAGCCAGGGTCTTTCTGCTGAGGACGAAACCTTCGGGAGAGGGCGACTGGAG 1904  
QY 1889 CAGTTCACGCTGTGGCAGCAAGGAGAAATGAAATGCTTCGCAAGAGAGCTCTTAA 1948  
Db 1905 CAGTTCACACTGTGGCAGCAAGGAGGAAATGAGAAATGCTGTGAAAGGAGCGCTAA 1964  
QY 1949 AACCTGTACCTTACTAGAAAAGTTCCCGAGACACAGGATGAGAGAGGAGCAGATCAA 2008  
Db 1965 GACCTGTGCTTACTAGAAAAGTTTCCGAAACACAGGATGAGAGAGGAGCAGATCAA 2024  
QY 2009 ATATTCCATCATGCCCGGAGCTCACTGTGGCCGACACAGGCGCCACAACTGCGAG 2068  
Db 2025 ATACTCCATCATGCCCTGGAACCTCATGCTGTGGCCGATACAGGACCCACAACTGCGAG 2084  
QY 2069 GCTCGAATGCACTGGGCTTGGTATTCCTCCAGAGAGGCTGCAAGATTGCGATGCGCAA 2128  
Db 2085 GCTCGAATGCACTGGGCTTGGTATTCCTCCAGAGAGGCTGCAAGATTGCGATGCGCAA 2144  
QY 2129 CGAGACCAG 2137  
Db 2145 TGAGACCAG 2153

## RESULT 7

BM806632 1050 bp mRNA linear EST 05-MAR-2002  
LOCUS AGENCOURT 6543030 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5548943  
DEFINITION 5', mRNA sequence.

ACCESSION BM806632.1 GI:19123455

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LLAM12258 row: k column: 24

High quality sequence stop: 649.

Location/Qualifiers

1. 1050

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5548943"

/tissue\_type="melanotic melanoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_72"

/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

## ORIGIN

Query Match 39.0%; Score 905.2; DB 4; Length 1050;  
Best Local Similarity 97.6%; Pred. No. 6.8e-221;  
Matches 961; Conservative 0; Mismatches 19; Indels 5; Gaps 4;  
QY 865 ATATCTCTGTAGAGATTACAGAGTAATTTGTAGAGAGTAAGCATTTTCTCTGGAG 924  
Db 1 ATATCTCTGTAGAGATTACAGAGTAATTTGTAGAGAGTAAGCATTTTCTCTGGAG 60

QY 925 AACAGCAGGAGTACCAACCCAGAAAAAAATAGAAAAACAGATGATCCAGAACAAAAAGCAA 984  
Db 61 AACAGCAGGAGTACCAACCCAGAAAAAAATAGAAAAACAGATGATCCAGAACAAAAAGCAA 120  
QY 985 AAGTTAAGAAAAAGAACCCCTAAACCTTTTAAATTAATTTGATTAAGACTATTAAAGCTGGAAC 1044  
Db 121 AAGTTAAGAAAAAGAACCCCTAAACCTTTTAAATTAATTTGATTAAGACTATTAAAGCTGGAAC 180  
QY 1045 TTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAATTTGAGGAGCAGTGAATGCAATTTA 1104  
Db 181 TTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAATTTGAGGAGCAGTGAATGCAATTTA 240  
QY 1105 AAGAACTAGTACGCAAAATACCTCAGAGTCCACGAGCAAGATATGGGAAGGGCGAGTGTG 1164  
Db 241 AAGAACTAGTACGCAAAATACCTCAGAGTCCACGAGCAAGATATGGGAAGGGCGAGTGTG 300  
QY 1165 AGGATGATTTGGCTGAGAAAGAGGAGAGTAATGAGGTGCTACGTTGGAGCCATCGAGACCT 1224  
Db 301 AGGATGATTTGGCTGAGAAAGAGGAGAGTAATGAGGTGCTACGTTGGAGCCATCGAGACCT 360  
QY 1225 ACCAAGAGGTGGCCAGCCTACCTGATGTCCTCGAGACCTGCTGAAAGCTGAGTTTGAAGC 1284  
Db 361 ACCAAGAGGTGGCCAGCCTACCTGATGTCCTCGAGACCTGCTGAAAGCTGAGTTTGAAGC 420  
QY 1285 GTGCTCAGACAGGCAACAAATTTCTAGGTGATATGAGAGGTTCCTGCTTACCTGCGAGA 1344  
Db 421 GTGCTCAGACAGGCAACAAATTTCTAGGTGATATGAGAGGTTCCTGCTTACCTGCGAGA 480  
QY 1345 GATTAGTTCAATTTTCCCAATGATCTCTTAAAAAATGACTTGGCGTGGGATACC 1404  
Db 481 GATTAGTTCAATTTTCCCAATGATCTCTTAAAAAATGACTTGGCGTGGGATACC 540  
QY 1405 TCTTGATGAGAGATATGACAAATGCAAGAAAGTTTATGAAGAGGTGCTGAGTGTGACAC 1464  
Db 541 TCTTGATGAGAGATATGACAAATGCAAGAAAGTTTATGAAGAGGTGCTGAGTGTGACAC 600  
QY 1465 CTATGATGCTTTGCTAAAGTCCATTTATGAGTCTTATGAGAGGTGCTGAGTGTGACAC 1524  
Db 601 CTATGATGCTTTGCTAAAGTCCATTTATGAGTCTTATGAGAGGTGCTGAGTGTGACAC 660  
QY 1525 CTGAGAGCATCCCATATTTAAAGAGGATAGAAATCGGAGATCTTGGCACTGATGATG 1584  
Db 661 CTGAGAGCATCCCATATTTAAAGAGGATAGAAATCGGAGATCTTGGCACTGATGATG 720  
QY 1585 GGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGAGGTGGGAAACAAAGAGGCATATA 1644  
Db 721 GGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGAGGTGGGAAACAAAGAGGCATATA 780  
QY 1645 AGTGGTATGAGCTTTGGGCACAAAGAGGAGACACTTTTGATCTGCTGCGCAACCTCACTCT 1704  
Db 781 AGTGGTATGAGCTTTGGGCACAAAGAGGAGACACTTTTGATCTGCTGCGCAACCTCACTCT 840  
QY 1705 ACAATGTGATGAGCTTTGGAAGCAGACGCTTGGTGGACCCCAAGAAAGAGGCGGTACACAG 1764  
Db 841 ACAATGTG-ATGAGCTGGAAGCAGACCTTTGGTGGACCCCAAGAAAGAGGCGGTACACAG 899  
QY 1765 AGTTAGTAAAGT-CTTTAGAAAGAACTGGAAGTTAAT-CCGAGATGAAGGCC--TTGCA 1820  
Db 900 AGTTAGTAAAGTCTTTTGAAGAAAGAACTGGAAGTTAATCCCGAAATGAAGGCCCTGTGCA 959  
QY 1821 GTGATGATAAGCCAAAGGCTCTCT 1845  
Db 960 GTGATGATAAGCCAAAGGCGCT 984

## RESULT 8

BM806636  
LOCUS BX403636 1011 bp mRNA linear EST 01-MAY-2004  
DEFINITION BX403636 Homo sapiens PLACENTA Homo sapiens cDNA clone CL08A0072C04  
S-PRIME, mRNA sequence.  
ACCESSION BX403636  
VERSION BX403636.2 GI:46924400

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1011)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30635043.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4537.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CL0BA007ZC04RPl&c=4537.r.
FEATURES
Location/Qualifiers
1..1011
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BA007ZC04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 37.6%; Score 873.4; DB 5; Length 1011;
Best Local Similarity 97.6%; Pred. No. 9.9e-213;
Matches 880; Conservative 5; Mismatches 17; Indels 0; Gaps 0;
QY 1 CGGACCGGTGCAATGGCCAGCGTGAAGATGCCAGAGCAGCGGCAACAGCAGCAGCAGCG 60
DB 110 CGGACCGGTGCAATGGCCAGCGTGAAGATGCCAGAGCAGCGGCAACAGCAGCAGCAGCG 169
QY 61 GCTCGGCGGTAGCAGTGGCGCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 120
DB 170 GCTCGGCGGTAGCAGTGGCGCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 229
QY 121 AGCATGGAGGACACAAGAAATGGGAGGAAGCGGAGCTCTCGGGAACCTTCATTTTCACGT 180
DB 230 AGCATGGAGGACACAAGAAATGGGAGGAAGCGGAGCTCTCAGGAACCTTCATTTTCACGT 289
QY 181 GGTTCATGGTATTCATTCGTGGCGCTCGACATCTGTAGCTGTGCTGGTTGGTTGATC 240
DB 290 GGTTCATGGTATTCATTCGTGGCGCTCGACATCTGTAGCTGTGCTGGTTGGTTGATC 349
QY 241 TTGTTGACTATGAGGAAGTTCTAGAAAACCTAGGAATCTATCATGCTGATGGTATGGAG 300
DB 350 TTGTTGACTATGAGGAAGTTCTAGAAAACCTAGGAATCTATCATGCTGATGGTATGGAG 409
QY 301 ATTTTGTATGGTATGCCAAGTTTATTAGGACTTTAAAGAGAGATCTACTTTCAGAGC 360
DB 410 ATTTTGTATGGTATGCCAAGTTTATTAGGACTTTAAAGAGAGATCTACTTTCAGAGC 469
QY 361 CAGCAGTCCCGCAGAGAGCTGACCACTAGCAGCCCGGAGGAGCAGGTTCTCTGGG 420
DB 470 CAGCAGTCCCGCAGAGAGCTGACCACTAGCAGCCCGGAGGAGCAGGTTCTCTGGG 529
QY 421 AGGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAATTCAGTCCCTTCTCCATG 480
DB 530 AGGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAATTCAGTCCCTTCTCCATG 589

```

---

```

QY 481 AAATGGTACACCGCAGAAACATGTTGAGGGAGAGAGACTTGCACCAAGAGAGATGGACCCACAG 540
DB 590 AAATGNTACACCGCAGAAACATGTTGAGGGAGAGAGACTTGCACCAAGAGAGATGGACCCACAG 649
QY 541 GAGAACCAACCAAGAGAGATGATGATGTTCTTATGGCGACTGATGTAGATGATGATTTG 600
DB 650 GAGAACCAACCAAGAGAGATGATGATGTTCTTATGGCGACTGATGTAGATGATGATTTG 709
QY 601 AGACCCCTGGAACCTGAAAGTATCTCATGAAGAACCGAGCATAGTTACCAACGTTGGAAGAGA 660
DB 710 AGACCCCTGGAACCTGAAAGTATCTCATGAAGAACCGAGCATAGTTACCAACGTTGGAAGAGA 769
QY 661 CAGTTTCAACAGCTGATATCAGGATATGGAAGAGATGATGTCTGAGCAGGAGAAATCCAG 720
DB 770 CAGTTTCAACAGCTGATATCAGGATATGGAAGAGATGATGTCTGAGCAGGAGAAATCCAG 829
QY 721 ATTCCAGTGAACAGTAGTAGAAGATGGAAGATGACCATGATACAGATGATGTAACAT 780
DB 830 ATTCCAGTGAACAGTAGTAGAAGATGGAAGATGACCATGATACAGATGATGTAACAT 889
QY 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840
DB 890 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 949
QY 841 CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAAGATTCACAGGTAATTTGTAGAAG 900
DB 950 CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAAGATTCACAGGTAATTTGTAGAAG 1009
QY 901 AA 902
DB 1010 TA 1011
RESULT 9
BX459083
LOCUS
DEFINITION BX459083 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YB13
5-PRIME, mRNA sequence.
ACCESSION BX459083
VERSION BX459083.2 GI:47064816
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1085)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 22, 2003 this sequence version replaced gi:31027087.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4537.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DE012CA07QPl&c=4537.r.
FEATURES
Location/Qualifiers
1..1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YB13"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into

```

the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN

Query Match 35.9%; Score 834.6; DB 5; Length 1085;  
Best Local Similarity 99.3%; Pred. No. 9.3e-203;  
Matches 837; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 113 AGAGCAAAAGCATGAGGACACAAGAATGGAGGAAAGCGGACTCTCGGAACTTCATT 172  
DB 1 AGAGCAAAAGCATGAGGACACAAGAATGGAGGAAAGCGGACTCTCGGAACTTCATT 60

QY 173 CTTACGCTGTTTATGTTGATGATTCATTGCTGGGCTCTGACATCTCTAGCTCTCGTTTG 232  
DB 61 CTTACGCTGTTTATGTTGATGATTCATTGCTGGGCTCTGACATCTCTAGCTCTCGTTTG 120

QY 233 GTTTGATCTTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGG 292  
DB 121 GTTTGATCTTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGG 180

QY 293 TGATGGAGATTTGATGTTGATGATGCCAAAGTTTTTATAGGACTTAAAGAGAGATCTAC 352  
DB 181 TGATGGAGATTTGATGTTGATGATGCCAAAGTTTTTATAGGACTTAAAGAGAGATCTAC 240

QY 353 TTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGT 412  
DB 241 TTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGT 300

QY 413 TCCTGTGGAGGAGAACCCAGAAATATCGAGATGAGCAAAAGAAACAAATTCAGTCCCT 472  
DB 301 TCCTGTGGAGGAGAACCCAGAAATATCGAGATGAGCAAAAGAAACAAATTCAGTCCCT 360

QY 473 TCTCATGAATGTTGATGAGGAGAGTGTGAGGAGAGACTTGCACAGAGATGG 532  
DB 361 TCTCATGAATGTTGATGAGGAGAGTGTGAGGAGAGACTTGCACAGAGATGG 420

QY 533 ACCACAGGAGAACCCACAGAGAGATGATGATTTCTTATGGCGACTGATGATGA 592  
DB 421 ACCACAGGAGAACCCACAGAGAGATGATGATTTCTTATGGCGACTGATGATGA 480

QY 593 TAGATTTGAGACCTTGGAACTGAAATATCTCATGAAGAAACCGAGCATAGTTACCACT 652  
DB 481 TAGATTTGAGACCTTGGAACTGAAATATCTCATGAAGAAACCGAGCATAGTTACCACT 540

QY 653 GGAAGACAGATTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTTGAGCAGGA 712  
DB 541 GGAAGACAGATTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTTGAGCAGGA 600

QY 713 AAATCCAGATTCAGTGAACAGTAGTAGAGATGAAGATTTGCCATGATACAGATGA 772  
DB 601 AAATCCAGATTCAGTGAACAGTAGTAGAGATGAAGATTTGCCATGATACAGATGA 660

QY 773 TGTAACATACCAAGTCTATGAGGAAACAAGCAGTATATGAACCTCTAGAAAATGAAGGAT 832  
DB 661 TGTAACATACCAAGTCTATGAGGAAACAAGCAGTATATGAACCTCTAGAAAATGAAGGAT 720

QY 833 AGAAATCACAAGAGTAACCTGCTCCCTGAGATAATCTCTGTAAGATTTACAGCTAAT 892  
DB 721 AGAAATCACAAGAGTAACCTGCTCCCTGAGATAATCTCTGTAAGATTTACAGCTAAT 780

QY 893 TGTAAGAGAGTAAGCATTTTCTGTTGAGAGACAGAGAGATGACCAAGAACAAA 952  
DB 781 TGTAAGAGAGTAAGCATTTTCTGTTGAGAGACAGAGAGATGACCAAGAACAAA 840

QY 953 TAG 955  
DB 841 AAG 843

RESULT 10  
AUI30952  
LOCUS  
DEFINITION AUI30952 NT2RP3 Homo sapiens cDNA clone NT2RP3001731 5', mRNA  
859 bp mRNA linear EST 01-AUG-2002

sequence.  
AUI30952 GI:10991306  
EST.  
Homo sapiens (human)  
Homo sapiens  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,  
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and  
Isogai, T.  
HRI human cDNA project  
Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.  
Location/Qualifiers  
1. 859  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xrefs="taxon:9606"  
/clone="NT2RP3001731"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2P3"  
/notes="Vector: pME189FL3; mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"

ORIGIN  
Query Match 33.9%; Score 787.6; DB 1; Length 859;  
Best Local Similarity 96.5%; Pred. No. 1e-190;  
Matches 834; Conservative 0; Mismatches 25; Indels 5; Gaps 3;

QY 1013 AAATAAATTTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAACCTCCGTAAGAG 1072  
DB 1 AAATAAATTTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAACCTCCGTAAGAG 60

QY 1073 GGGAAAATTTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTTCAGAG 1132  
DB 61 GGGAAAATTTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTTCAGAG 120

QY 1133 TCACACGAGCAAGATATATGGGAAGCGCAGTGTGAGGATGATTGGCTGAGAAGAGGAGAAG 1192  
DB 121 TCACACGAGCAAGATATATGGGAAGCGCAGTGTGAGGATGATTGGCTGAGAAGAGGAGAAG 180

QY 1193 TAATGAGTGTCTACGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCTACCTGATGT 1252  
DB 181 TAATGAGTGTCTACGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCTACCTGATGT 240

QY 1253 CCCTGCGAGCCTCTGAAGCTGAGTTGAACGCTCGCTCAGACAGGCAAAATTTCTAGG 1312  
DB 241 CCCTGCGAGCCTCTGAAGCTGAGTTGAACGCTCGCTCAGACAGGCAAAATTTCTAGG 300

QY 1313 TCATATGAGAGGTTCCCTGCTTTACCTTCAGAGATTAAGTTCAACTATTTTCCCAATGATAC 1372  
DB 301 TCATATGAGAGGTTCCCTGCTTTACCTTCAGAGATTAAGTTCAACTATTTTCCCAATGATAC 360

QY 1373 TTCCTTAAAAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCAAAA 1432  
DB 361 TTCCTTAAAAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCAAAA 420

QY 1433 GAAAGTTTATGAGAGGTTCTGAGTGTGACACCTTAATGATGGCTTTTCTTAAGTCCATTA 1492  
DB 421 GAAAGTTTATGAGAGGTTCTGAGTGTGACACCTTAATGATGGCTTTTCTTAAGTCCATTA 480

```
QY 1493 TGGCTTCATCTGTAAGGCACAGACAAAATGCTGAGAGCATCCCATATTTAAAGGAAGG 1552
Db 481 TGGCTTCATCTGTAAGGCACAGACAAAATGCTGAGAGCATCCCATATTTAAAGGAAGG 540
QY 1553 AATAGAATCCGGAGATCCTGGCACTGATGATGGAGAGATTTATTTCCACCTGGGGGATGC 1612
Db 541 AATAGAATCCGGAGATCCTGGCACTGATGATGGAGAGATTTATTTCCACCTGGGGGATGC 600
QY 1613 CATCAGAGGTTGGGAACAAGAGGATATTAAGTGTATGAGCTTGGGCACAGAGAGG 1672
Db 601 CATCAGAGGTTGGGAACAAGAGGATATTAAGTGTATGAGCTTGGGCACAGAGAGG 660
QY 1673 ACATTTTGATCTCTCTGGCAACCTCACTCTCAATGTGAATGCACTGAAAGCACAGCC 1732
Db 661 ACATTTTGATCTCTCTGGCAACCTCACTCTCAATGTGAATGCACTGAAAGCACANCC 720
QY 1733 TTGCTGACCCCAAAAGAAAGCGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTG 1792
Db 721 TTGCTGACCCCAAAAGAAAGCGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTG 779
QY 1793 GAAGTTAATCCAGATGAAGCCCTGACGATGATGAAGCCAAAGGTCTCTTCCTGCC 1852
Db 780 GAAGTTAATCCAGATGAAGCCCT--GCATGATGGATAA--GCCAAGGTCTNTTTCTGNC 835
QY 1853 TGAGATGAACAACTGAGGGGAAAA 1876
Db 836 TGAGATGAACAACTGAGGGGAAAA 859
```

```
RESULT 11
BU838423
LOCUS 930 bp mRNA linear EST 16-OCT-2002
DEFINITION AGNCOURT_8123161 Lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6177038 5', mRNA sequence.
ACCESSION BU838423
VERSION BU838423.1 GI:24022818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 930)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13554 row: f column: 15
High quality sequence start: 66
High quality sequence stop: 549.
Location/Qualifiers
1..930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177038"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGGGTCGCG-3' and
```

## FEATURES

source

```
5'-GACTAGTTCTAGATCCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
ORIGIN
Query Match 31.0%; Score 719.8; DB 5; Length 930;
Best Local Similarity 96.3%; Pred. No. 2.7e-173;
Matches 758; Conservative 0; Mismatches 27; Indels 2; Gaps 2;
QY 1 CGACCGTGCATATGCGCCAGCGTAAGATGCGCAAGAGCAGCGCAACAGCAGCAGCAGCG 60
Db 138 CGACCGTGCATATGCGCCAGCGTAAGATGCGCAAGAGCAGCGCAACAGCAGCAGCAGCG 197
QY 61 GCTCCGCGCAGCGGTAGCAGCGTGGCGCAGCAGCGCGCGCGCGCGCGCGGAGAGACAA 120
Db 198 GCTCCGCGCAGCGGTAGCAGCGTGGCGCAGCAGCGCGCGCGCGCGCGGAGAGACAA 257
QY 121 AGCATGGAGGACACAAAGAATGGGAGGAAAGCGGACTCTCGGGAACCTTCAATTTTCACGT 180
Db 258 AGCATGGAGGACACAAAGAATGGGAGGAAAGCGGACTCTCAGGAACCTTCAATTTTCACGT 317
QY 181 GGTATTATGTTGATGTCCTGGCGGTCTGGACATCTGTAGCTGCTGTTGTTGATC 240
Db 318 GGTATTATGTTGATGTCCTGGCGGTCTGGACATCTGTAGCTGCTGTTGTTGATC 377
QY 241 TTCTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGGAG 300
Db 378 TTGTGTACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGGAG 437
QY 301 ATTTTGATGTGGATGATGCCAAAGTTTTATTTAGGACTTAAAGAGAGATCTACTTCAGAGC 360
Db 438 ATTTTGATGTGGATGATGCCAAAGTTTTATTTAGGACTTAAAGAGAGATCTACTTCAGAGC 497
QY 361 CAGCAGTCCCGCAGAAAGAGGCTGACCAACACTGAGCCCGCAGGAGCAGGTTCTCTGTCG 420
Db 498 CAGCAGTCCCGCAGAAAGAGGCTGACCAACACTGAGCCCGCAGGAGCAGGTTCTCTGTCG 557
QY 421 AGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480
Db 558 AGCAGAACCCCAAGATATCGAAGATGAAGCAAGAGAGCAAAATTCAGTCCCTTCTCCATG 617
QY 481 AAATGGTACACGACAGCAACATGTTGAGGAGAGAGACTTGCACCAAGAGATGGA-CCACA 539
Db 618 AAATGGTACACGACAGCAACATGTTGAGGAGAGAGACTTGCACCAAGAGATGGA-CCACA 677
QY 540 GGAGAACCAACACAGAGGATGATGAGTTCTTTATGCGGACTGATGTAGATGATGATTT 599
Db 678 GGAGAACCAACACAGAGGATGATGAGTTCTTTATGCGGACTGATGTAGATGATGATTT 737
QY 600 GAGACCTTGGAACTGTAAGTATCTCATGAAGAACCCGAGCATAGTTACACCGTGGAAAGAG 659
Db 738 GAGACCTTGGCACTCTGAAGTATCTCATGAAGAACCCGAGCATAGTTACACCGGAAAGAA 797
QY 660 ACAGTTTTCACAGACTGTAAATCAGGATATCGAAGAGATGATGTGAGCAGGAGAAATCCA 719
Db 798 ACAGTTTTCACAGACTGTAAATCAGGATATCGAAGAGATGATGTGAGCAGGAGAAATCCA 857
QY 720 GATTCCAGTGAACCAAGTAGTAGAGATG-AAAGTTTGCACCATGATACAGATGATGTAAC 778
Db 858 GATTCCAGTGTAGCCCTTATTAGAAGATGAAAGATGCCCCCATATATACCGATGTTTAC 917
QY 779 ATACCAA 785
Db 918 ATATCCA 924
```

RESULT 12

CK724966/c

LOCUS

DEFINITION

782 bp mRNA linear EST 05-AUG-2004  
UI-H-BD1-axp-c-22-0-UI.s1 NCI\_CGAP\_ED1 Homo sapiens cDNA clone



```

ACCESSION      UI-H-EDI-axp-c-22-0-UI 3', mRNA sequence.
VERSION         CK724966
KEYWORDS        1  GI:42535838
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1  (bases 1 to 782)
AUTHORS         NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT         Contact: Daniela S. Gerhard, Ph.D.
                Office of Cancer Genomics
                National Cancer Institute
                Bldg: 31 Rm10A07 Bethesda, MD 20892
                Email: cgapbs-re@mail.nih.gov
                Tissue Procurement: Dr. Jose Mercuende
                cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Distribution information can be found at
                http://genome.uiowa.edu/distribution/cgap.html
                Seq primer: M13 FORWARD
                POLYA=Yes.

FEATURES             Location/Qualifiers
     source           1..782
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-EDI-axp-c-22-0-UI"
                     /tissue_type="Chondrosarcoma"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI CGAP EDI"
                     /notes="Organ: Left Pubic Bone; Vector: pTV73-Pac
                     (Pharmacia) with a modified polylinker; Site 1: EcoR I;
                     Site 2: Not I; NCI CGAP EDI is a normalized cDNA library
                     containing the following tissue(s): Chondrosarcoma cell
                     line C85. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pTV73-pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is GCTCAAGGCT.
                     TAG TISSUE=chondrosarcoma
                     TAG LIB=UI-H-EDI
                     TAG_SEQ=CCTCAAGGCT"

ORIGIN
Query Match      30.0%; Score 697; DB 7; Length 782;
Best Local Similarity 98.9%; Pred. No. 1.9e-167;
Matches 709; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1210 GAGCCATCGAGACCTACCAAGAGGTGGCCAGCGCTACCTGATGTCCTCGAGACCTGGCTGA 1269
      |||||
DB 716 GAGCCATCGAGACNTACCAAGA-GTGGCCAGCGCTACCTGATGTCCTCGAGACCTGGCTGA 658
      |||||
QY 1270 AGCTGAGTTTGAAGGCTGCTCAGACAGGCAACAATTTCTAGTGCATATGAGAGTTCCC 1329
      |||||
DB 657 AGCTGAGTTTGAAGGCTGCTCAGACAGGCAACAATTTTAGTGCATATGAGAGTTCCC 598
      |||||
QY 1330 TGCTTACCCTGCAGAGATAGTTCACTATTTCCTCAATGATATCTTCTTAAAAAATGACC 1389
      |||||
DB 597 TGCTTACCCTGCAGAGATAGTTCACTATTTCCTCAATGATATCTTCTTAAAAAATGACC 538
      |||||
QY 1390 TTGGCGTGGGATACCTCTTGATAGGAGATAATGACAATCAAGAGAAAGTTTATCAAGAGG 1449
      |||||

```

```

537 TTGGCGTGGGATACCTCTTGATAGGAGATAATGACAATGCAAGAAAGTTTATGAAGAGG 478
1450 TGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTTCATCTGAAGG 1509
      |||||
477 TGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTTCATCTGAAGG 418
1510 CACGAACAAAATTTGCTGTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAATCCGGAGATC 1569
      |||||
417 CACGAACAAAATTTGCTGTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAATCCGGAGATC 358
1570 CTGGCACTGATGTGGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGAGGGTTGGGA 1629
      |||||
357 CTGGCACTGATGTGGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGAGGGTTGGGA 298
1630 ACAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGGACACTTTTGCATCTGTCT 1689
297 ACAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGGACACTTTTGCATCTGTCT 238
1690 GGCACCTCTCACTCTCAATGTGAATGCACTGAAAGCACAGCCCTTGGTGACCCCAAAAG 1749
237 GGCACCTCTCACTCTCAATGTGAATGCACTGAAAGCACAGCCCTTGGTGACCCCAAAAG 178
1750 AAACGGGCTACACAGAGATTAGTAAAGTCTTTAGAAAGAAACTCGAAAGTTAAATCCGAGATG 1809
      |||||
177 AAACGGGCTACACAGAGATTAGTAAAGTCTTTAGAAAGAAACTCGAAAGTTAAATCCGAGATG 118
1810 AAGGCTTTCAGTGTGATGATTAAGCCAAAGTCTCTTCTGCTGAGGATGAAACCTGA 1869
      |||||
117 AAGGCTTTCAGTGTGATGATTAAGCCAAAGTCTCTTCTGCTGAGGATGAAACCTGA 58
1870 GGGAAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAAATGAAA 1926
      |||||
57 GGGAAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAAATGAAA 1
      |||||

RESULT 13
BM790814      689 bp mRNA linear EST 05-MAR-2002
LOCUS        K-EST0070696 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-16-C08 5',
DEFINITION   mRNA sequence.
ACCESSION    BM790814
VERSION      BM790814.1 GI:19139046
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 689)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 16 row: C column: 08
High quality sequence stop: 689.
Location/Qualifiers
1..689
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-16-C08"
/sex="F"
/tissue_type="Stomach"
/cell_type="floating aggregates"
/lab_host="SNU520"
/lab_host="Top10f"

```





Mon Mar 28 06:07:00 2005

Qy	944	AGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCC	1003
Db	552	AGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCC	611
Qy	1004	TAAACTTTTAAATATAATTTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAAAACT	1063
Db	612	TAAACTTTTAAATATAATTTGATAAGACTATTAAAGCTGAACTTGATGCTGCAG-AAAACT	670
Qy	1064	CCGTAAAAAGGGGCAAAATTTGAGGAGCAGTGAAT	1097
Db	671	CCGTAAAGGGGCAAAATTTGAGGAGCAGTGAAT	704

Search completed: March 25, 2005, 14:28:59  
Job time : 6791 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 24, 2005, 12:41:52 ; Search time 1157 Seconds  
(without alignments)  
11890.647 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatgcccgag.....cttgggaactctggagaga 2324

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Genesegm16Dec04:\*  
1: Genesegm1980s:\*  
2: Genesegm1990s:\*  
3: Genesegm2000s:\*  
4: Genesegm2001s:\*  
5: Genesegm2001bs:\*  
6: Genesegm2002as:\*  
7: Genesegm2002bs:\*  
8: Genesegm2003as:\*  
9: Genesegm2003bs:\*  
10: Genesegm2003cs:\*  
11: Genesegm2003ds:\*  
12: Genesegm2004as:\*  
13: Genesegm2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	4	Aaf89811 cDNA enco
2	2324	100.0	2324	8	ABs57868 Human cDN
3	2324	100.0	2324	9	ADA00640 cDNA enco
4	2320.8	99.9	2452	13	ADP23914 PRO polyp
5	2320.8	99.9	2452	13	ADP23914 PRO polyp
6	2319.2	99.8	5358	10	ADJ56257 Bovine cD
7	947	40.7	2680	6	ABX04178 Human mRN
8	947	40.7	2680	13	ADP24208 PRO polyp
9	819	35.2	2442	2	Aaz23609 Human lab
10	819	35.2	2442	6	ABK39743 cDNA enco
11	819	35.2	2442	8	ACA12072 Human lun
12	819	35.2	2442	8	ACA03258 lung canc
13	819	35.2	2442	10	ADH47310 Human lun
14	819	35.2	2442	13	ADJ21229 Human lun
15	711.4	30.6	2648	10	ACA92448 Human lun
16	617.6	26.6	3110	11	ACN88788 Breast ca
17	464.6	20.0	502	9	ACH47067 Human inf
18	416.2	17.9	660	6	ABQ59507 Human col
19	317	13.6	495	11	ADT94758 Colon can
20	213.6	9.2	1662	4	ABL10339 Drosophil

21	170.6	7.3	1382	2	AAx22124	Human sec
22	170.6	7.3	1382	8	ADA39640	Human sec
23	170.6	7.3	1382	8	ACC50349	Human sec
24	170.6	7.3	1382	10	ADA55835	Gene enco
25	153.2	6.6	161	12	ACH93769	Human gen
26	153.2	6.6	592	12	ACH80069	Human gen
27	147.6	6.4	3986	4	ABL10338	Drosophil
28	127.4	5.5	583	6	ABQ59571	Human col
29	113.4	4.9	1686	2	Aaq87587	DNA enco
30	112.2	4.8	3399	2	AAT05868	Chicken 1
31	103.2	4.4	3755	4	ABL10352	Drosophil
32	96.8	4.2	671	3	AAC95033	Cat flea
33	87.8	3.8	3579	3	AAa70099	Plasmodiu
34	81.8	3.5	4590	1	AAa60472	Sequence
35	74.8	3.2	1113	12	ADQ00111	Novel hum
36	74.6	3.2	3095	2	ADN98542	Sequence
37	74.6	3.2	17137	6	AAQ03875	Novel hum
38	72.2	3.1	31124	11	ACN44156	Mouse imm
39	71.2	3.1	3912	8	ACA22058	Prokaryot
40	70.4	3.0	15507	8	AAL53680	Genomic D
41	70.2	3.0	1997	13	ADR07778	Full leng
42	68.6	3.0	585	2	AAQ87589	DNA enco
43	68.6	3.0	996	2	AAQ87588	DNA enco
44	67.4	2.9	3337	2	AAT34620	P. vivax
45	67.4	2.9	3337	2	AAT34620	P. vivax

## ALIGNMENTS

## RESULT 1

AAF89811

ID AAF89811 standard; cDNA; 2324 BP.

AC AAF89811;

DT 23-JUL-2001 (first entry)

DE cDNA encoding human aspartyl (asparaginyl) beta-hydroxylase (HAAH).

EW Epidermal growth factor-like domain; EGF-like domain; cancer;

KW human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour; ss.

OS Homo sapiens.

PH Key

FT CDS

FT Location/Qualifiers

FT 12..2289

FT /\*tag= a

FT /product= "human aspartyl (asparaginyl) beta-hydroxylase"

PN WO200135102-A2.

PP 17-MAY-2001.

PD 08-NOV-2000; 2000WO-US030738.

PF 08-NOV-1999; 99US-00436184.

PR (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.

PA Wands JR, De La Monte SM, Ince N, Carlson RI;

PP 2001-329171/34.

DR P-FSDS; AAB83919.

XX Diagnosing malignant neoplasm in a mammal, involves contacting mammalian

XX sample with antibody that binds to human aspartyl beta-hydroxylase

XX polypeptide to form antigen-antibody complex and detecting the complex.

PS Disclosure; Page 6-7; 76pp; English.

XX The present sequence encodes a human aspartyl (asparaginyl) beta-

XX hydroxylase (HAAH) enzyme. Epidermal growth factor (EGF)-like domains of

CC polypeptides are hydroxylated by HAAH enzymes. HAAH is used in the method  
CC of the invention. The specification describes a method for diagnosing a  
CC malignant neoplasm in a mammal. The method comprises contacting a body  
CC fluid with an antibody which binds to HAAH polypeptide under complex  
CC forming conditions, and detecting the antigen-antibody complex. The  
CC method is useful for diagnosing and prognosing a malignant neoplasm in a  
CC bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
CC serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
CC where the neoplasm is derived from endodermal tissue and is selected from  
CC colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
CC bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
CC agent, are useful for killing tumour cells

XX  
SQ Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;

Query Match 100.0%; Score 2324; DB 4; Length 2324;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGGACCTGCAATGGCCAGCGTAAGAAATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG	60
DB	1	CGGACCTGCAATGGCCAGCGTAAGAAATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG	60
QY	61	GCTCCGCGAGCGGTAGCACAGAGTCGGGCGAGCAGCAGCCCGCGGCGCCGGAGAGACAA	120
DB	61	GCTCCGCGAGCGGTAGCACAGAGTCGGGCGAGCAGCAGCCCGCGGCGCCGGAGAGACAA	120
QY	121	AGCATGAGGAGACACAAGATGGGAGAAAGCGGACCTCTCGGNACTTCATCTTCACGT	180
DB	121	AGCATGAGGAGACACAAGATGGGAGAAAGCGGACCTCTCGGNACTTCATCTTCACGT	180
QY	181	GGTTATGCTGATTCGATTCGTCGGCGTCGGACATCTGTAGCTGTCTTGGTTTGATC	240
DB	181	GGTTATGCTGATTCGATTCGTCGGCGTCGGACATCTGTAGCTGTCTTGGTTTGATC	240
QY	241	TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGTGATGGAG	300
DB	241	TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGTGATGGAG	300
QY	301	ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	360
DB	301	ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	360
QY	361	CAGCAGTCCCGCAGAGAGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG	420
DB	361	CAGCAGTCCCGCAGAGAGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG	420
QY	421	AGGCAGAACCCGAGAAATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCATG	480
DB	421	AGGCAGAACCCGAGAAATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCATG	480
QY	481	AAATGGTACACGAGAAATGTTGAGGAGAGAGCTTGCACACAGAGATGAGCCACAG	540
DB	481	AAATGGTACACGAGAAATGTTGAGGAGAGAGCTTGCACACAGAGATGAGCCACAG	540
QY	541	GAGAACCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	600
DB	541	GAGAACCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	600
QY	601	AGACCTTGGAACTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG	660
DB	601	AGACCTTGGAACTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG	660
QY	661	CAGTTTCAACAGCTGTAATCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG	720
DB	661	CAGTTTCAACAGCTGTAATCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG	720
QY	721	ATTTCAGTGAACAGCTAGTAGAAGATGAAAGATTCACCATGATGATGATGATGATG	780
DB	721	ATTTCAGTGAACAGCTAGTAGAAGATGAAAGATTCACCATGATGATGATGATGATG	780
QY	781	ACCAAGTCTATGAGGAACAAGCAGGTATATGAACTCTAGAAATGAAAGGATAGAAATCA	840
DB	781	ACCAAGTCTATGAGGAACAAGCAGGTATATGAACTCTAGAAATGAAAGGATAGAAATCA	840

DB	781	ACCAAGTCTATGAGGAACAAGCAGGTATATGAACTCTAGAAATGAAAGGATAGAAATCA	840
QY	841	CAGAAAGTAACCTCCCTCAGGATAATCTCTGTAGAAGATTCACAGGTAATTTAGTAG	900
DB	841	CAGAAAGTAACCTCCCTCAGGATAATCTCTGTAGAAGATTCACAGGTAATTTAGTAG	900
QY	901	AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAAGTACCACAGAAACAAATAGAAA	960
DB	901	AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAAGTACCACAGAAACAAATAGAAA	960
QY	961	CAGATGATCCAGAACAAAGAAAGTAAAGAAAGAAAGCTTAACTTTTAAATAAAT	1020
DB	961	CAGATGATCCAGAACAAAGAAAGTAAAGAAAGAAAGCTTAACTTTTAAATAAAT	1020
QY	1021	TTGATAAGACTATTAAGCTGAACTTGATGCTGCAGAAAAAACTCCGTAAAGGGGAAA	1080
DB	1021	TTGATAAGACTATTAAGCTGAACTTGATGCTGCAGAAAAAACTCCGTAAAGGGGAAA	1080
QY	1081	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGGCAATACCTCAGAGTCCACG	1140
DB	1081	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGGCAATACCTCAGAGTCCACG	1140
QY	1141	CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATAGG	1200
DB	1141	CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATAGG	1200
QY	1201	TGCTAGCTGGAGCCTACGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCG	1260
DB	1201	TGCTAGCTGGAGCCTACGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCG	1260
QY	1261	ACCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAAATTTCTAGGTCAT	1320
DB	1261	ACCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAAATTTCTAGGTCAT	1320
QY	1321	GAGTTTCCCTGCTTACCTGCGAGATTAAGTTCAACTATTTCCCAATGATCTTCTTAA	1380
DB	1321	GAGTTTCCCTGCTTACCTGCGAGATTAAGTTCAACTATTTCCCAATGATCTTCTTAA	1380
QY	1381	AAATGACCTTGGCGGGATACCTTGTATAGGAGATTAATGCAATGCAAGAAAGTTT	1440
DB	1381	AAATGACCTTGGCGGGATACCTTGTATAGGAGATTAATGCAATGCAAGAAAGTTT	1440
QY	1441	ATGAAGAGTGTGAGTGTGACACCTTAATGATGCTTAAAGTCCATTATGGCTTCA	1500
DB	1441	ATGAAGAGTGTGAGTGTGACACCTTAATGATGCTTAAAGTCCATTATGGCTTCA	1500
QY	1501	TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGAAAT	1560
DB	1501	TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGAAAT	1560
QY	1561	CCGAGATCTCGCATGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGAGA	1620
DB	1561	CCGAGATCTCGCATGATGATGGAGATTTTATTTTCCACCTGGGGGATGCCATGAGA	1620
QY	1621	GGGTTGGGAACAAGAGGCATATAAGTGTATGATGCTTGGGCAACAGAGAGCACATTG	1680
DB	1621	GGGTTGGGAACAAGAGGCATATAAGTGTATGATGCTTGGGCAACAGAGAGCACATTG	1680
QY	1681	CATCTGTCTGGCAACCGCTCACTCTCAATGTGAATGAGTCTGAAAGCACAGCCTTGG	1740
DB	1681	CATCTGTCTGGCAACCGCTCACTCTCAATGTGAATGAGTCTGAAAGCACAGCCTTGG	1740
QY	1741	CCCCAAAGAAACGGGCTACACAGATTAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
DB	1741	CCCCAAAGAAACGGGCTACACAGATTAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
QY	1801	TCCGAGATGAGGCTTGCAGTATGATGATAAGCAAGGCTCTTCTCCCTGCTGAGGATG	1860
DB	1801	TCCGAGATGAGGCTTGCAGTATGATGATAAGCAAGGCTCTTCTCCCTGCTGAGGATG	1860
QY	1861	AAAACTTGGGAAAAAGGGGACTTGGAGCCAGTTTCCGCTGTGGCAGCAAGGAAGAA	1920
DB	1861	AAAACTTGGGAAAAAGGGGACTTGGAGCCAGTTTCCGCTGTGGCAGCAAGGAAGAA	1920





Db 421 AGGCAGAACCCGACATATCGAAGATGACGAAAGACAAATTCAGTCCCTTCTCCATG 480  
 Qy 481 AAATGGTACGCGAGAACATGTTGAGGGAAGACATTTGCAACAAAGAGATGGACCCACAG 540  
 Db 481 AAATGGTACGCGAGAACATGTTGAGGGAAGACATTTGCAACAAAGAGATGGACCCACAG 540  
 Qy 541 GAGAACACAAACAGAGATGATGAGTTCTTATGGCGACTGATGTAGATGATAGATTG 600  
 Db 541 GAGAACACAAACAGAGATGATGAGTTCTTATGGCGACTGATGTAGATGATAGATTG 600  
 Qy 601 AGACCTGGAACTGAAATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAGAGA 660  
 Db 601 AGACCTGGAACTGAAATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAGAGA 660  
 Qy 661 CAGTTTCAAGACTGTAATCAGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 720  
 Db 661 CAGTTTCAAGACTGTAATCAGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 720  
 Qy 721 ATTCAGTGAACCACTAGTAGAATGAAAGATTGCAACCATGATACAGATGATGTAACAT 780  
 Db 721 ATTCAGTGAACCACTAGTAGAATGAAAGATTGCAACCATGATACAGATGATGTAACAT 780  
 Qy 781 ACCAAGTCTATGAGGAAACCAAGCAGTATATGAACCTCTAGAAATGAAAGGATGAATCA 840  
 Db 781 ACCAAGTCTATGAGGAAACCAAGCAGTATATGAACCTCTAGAAATGAAAGGATGAATCA 840  
 Qy 841 CAGAAATGCTCTCCCTCCTGAGGATATCTGTTAGAGATTACAGGTAATTTAGAG 900  
 Db 841 CAGAAATGCTCTCCCTCCTGAGGATATCTGTTAGAGATTACAGGTAATTTAGAG 900  
 Qy 901 AGTAAAGCATTTTCTCTGGAAGAAACAGCAGGAGTACCAACGAAACCAATAGAAAA 960  
 Db 901 AGTAAAGCATTTTCTCTGGAAGAAACAGCAGGAGTACCAACGAAACCAATAGAAAA 960  
 Qy 961 CAGATGATCCAGAACCAAAAGCAAAAGTTAAGAAAGCAAGCTTAATTAAT 1020  
 Db 961 CAGATGATCCAGAACCAAAAGCAAAAGTTAAGAAAGCAAGCTTAATTAAT 1020  
 Qy 1021 TTGATAAGCATTTAAGCTGAATGATGCTGCAAGAAATCTCGTAAAGGGGAAAA 1080  
 Db 1021 TTGATAAGCATTTAAGCTGAATGATGCTGCAAGAAATCTCGTAAAGGGGAAAA 1080  
 Qy 1081 TTGAGGAACGATGATGATTTAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAG 1140  
 Db 1081 TTGAGGAACGATGATGATTTAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAG 1140  
 Qy 1141 CAAGATATGGGAAGCGCAGTGTAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
 Db 1141 CAAGATATGGGAAGCGCAGTGTAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
 Qy 1201 TGCTAGTGGAGCCATCGAGACTACCAAGAGGTGGCCAGCTTACCTGATGTCCCTGAG 1260  
 Db 1201 TGCTAGTGGAGCCATCGAGACTACCAAGAGGTGGCCAGCTTACCTGATGTCCCTGAG 1260  
 Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGTCTGAGACAGGCAACATTTCTAGTGTATGTA 1320  
 Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGTCTGAGACAGGCAACATTTCTAGTGTATGTA 1320  
 Qy 1321 GAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCACTATTTCCCAATGATCTTCTTAA 1380  
 Db 1321 GAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCACTATTTCCCAATGATCTTCTTAA 1380  
 Qy 1381 AAAATGACCTTGGCTGGGATACCTCTGTATAGGAGATATGCAATGCAAGAAAGTTT 1440  
 Db 1381 AAAATGACCTTGGCTGGGATACCTCTGTATAGGAGATATGCAATGCAAGAAAGTTT 1440  
 Qy 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTTAGGCTTCA 1500  
 Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTTAGGCTTCA 1500  
 Qy 1501 TCCTGAGGCACAGAACAAATTTGCTGAGAGCATCCATATTTAAAGGAGGAATAGAT 1560  
 Db 1501 TCCTGAGGCACAGAACAAATTTGCTGAGAGCATCCATATTTAAAGGAGGAATAGAT 1560

Qy 1561 CCGGAGATCTGGCACTGATGATGGAGATTTATTTTCCACTGGGGATGCCATGCAGA 1620  
 Db 1561 CCGGAGATCTGGCACTGATGATGGAGATTTATTTTCCACTGGGGATGCCATGCAGA 1620  
 Qy 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATAGAGCTTTGGGCAAGAGAGACACTTTG 1680  
 Db 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATAGAGCTTTGGGCAAGAGAGACACTTTG 1680  
 Qy 1681 CATCTGTCTGGCAACGCTCACTCTAATGTGAATGGACTGAAAGCAGAGCCTTGGTGA 1740  
 Db 1681 CATCTGTCTGGCAACGCTCACTCTAATGTGAATGGACTGAAAGCAGAGCCTTGGTGA 1740  
 Qy 1741 CCCCAGAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
 Db 1741 CCCCAGAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
 Qy 1801 TCGAGATGAAGGCTTGGCAGTGATGATAAGCAAGGCTCTTCTGCTGCTGAGGATG 1860  
 Db 1801 TCGAGATGAAGGCTTGGCAGTGATGATAAGCAAGGCTCTTCTGCTGCTGAGGATG 1860  
 Qy 1861 AAAACCTGAGGAAAGGGGACTGGAGCCAGTTCAACGCTGTGGCAGCAAGGAAGAA 1920  
 Db 1861 AAAACCTGAGGAAAGGGGACTGGAGCCAGTTCAACGCTGTGGCAGCAAGGAAGAA 1920  
 Qy 1921 ATGAAAATGCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
 Db 1921 ATGAAAATGCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
 Qy 1981 CACAGAGTGCAGAGAGGACAGATCAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040  
 Db 1981 CACAGAGTGCAGAGAGGACAGATCAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040  
 Qy 2041 GGCGGCACACAGGCCCCACAACTGAGGCTCGAATGCACTGGGCTGGTGATTTCCCA 2100  
 Db 2041 GGCGGCACACAGGCCCCACAACTGAGGCTCGAATGCACTGGGCTGGTGATTTCCCA 2100  
 Qy 2101 AGGAAGCTGCAAGATTGATGTCGCAACAGAGACAGGAGCTGGGAGGAGCAAGGTGC 2160  
 Db 2101 AGGAAGCTGCAAGATTGATGTCGCAACAGAGACAGGAGCTGGGAGGAGCAAGGTGC 2160  
 Qy 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGGAGGATGCTCATCTTTCCGCTGA 2220  
 Db 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGGAGGATGCTCATCTTTCCGCTGA 2220  
 Qy 2221 TATTCACTGTTGATGTTGGCATCCGGAATCGACACAGAGAGAGCAGCCTTCCAG 2280  
 Db 2221 TATTCACTGTTGATGTTGGCATCCGGAATCGACACAGAGAGAGCAGCCTTCCAG 2280  
 Qy 2281 CAATTTAGCATGATTTCAATCAAGCTTGGGAAACTCTGGAGAGA 2324  
 Db 2281 CAATTTAGCATGATTTCAATCAAGCTTGGGAAACTCTGGAGAGA 2324

RESULT 3  
 ADA00640  
 ID ADA00640 standard; cdna; 2324 BP.  
 XX  
 AC ADA00640;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE cdna encoding human aspartyl (asparaginyl) beta-hydroxylase (HAAH).  
 XX  
 KW Tumour growth inhibition; human aspartyl (asparaginyl) beta-hydroxylase;  
 KW HAAH hydroxylation; NOTCH polypeptide;  
 KW epidermal growth factor-like repeat; EGF-like repeat; tumour cell;  
 KW malignant neoplasm; colon cancer; breast cancer; pancreatic cancer;  
 KW liver cancer; cancer of the bile duct; cancer the central nervous system;  
 KW CNS; cytostatic; gene; ss; human.  
 XX  
 OS Homo sapiens.  
 XX

PH Key Location/Qualifiers  
 FT 12..2288  
 FT /\*tag= a  
 FT /product= "HAAH"  
 XX US2003031670-A1.  
 XX 13-FEB-2003.  
 XX 08-NOV-1999; 99US-00436184.  
 XX 08-NOV-1999; 99US-00436184.  
 XX (WAND/) WANDS J R.  
 XX (DMON/) DE LA MONTE S M.  
 XX (INCE/) INCE N.  
 XX (CARL/) CARLSON R I.  
 XX Wands JR, De La Monte SM, Ince N, Carlson RI;  
 XX WPI: 2003-605701/57.  
 XX P-PSDB; ADA00639.  
 XX Inhibiting tumor growth or killing tumor cells (e.g. cancer of the colon,  
 XX breast, pancreatic, liver or the central nervous system), by  
 XX administering an inhibitor of the human aspartyl (asparaginyl) beta-  
 XX hydroxylase.  
 XX Disclosure; Page 3-4; 30pp; English.  
 XX The present invention relates to a method for inhibiting tumour growth in  
 XX a mammal. The method comprises administering to the mammal a compound,  
 XX which inhibits the expression or enzymatic activity of a human aspartyl  
 XX (asparaginyl) beta-hydroxylase (HAAH). The compound may inhibit HAAH  
 XX hydroxylation of a NOTCH polypeptide. In particular, the compound may  
 XX inhibit hydroxylation of an epidermal growth factor (EGF)-like repeat  
 XX sequence in a NOTCH polypeptide. The methods are useful for inhibiting  
 XX tumour growth or killing tumour cells, or for diagnosing or  
 XX prognosticating a malignant neoplasm. In particular, the tumour or  
 XX neoplasm is colon cancer, breast cancer, pancreatic cancer, liver cancer,  
 XX cancer of the bile ducts, or cancer or tumour of the central nervous  
 XX system (CNS). The present sequence encodes HAAH.  
 XX Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;  
 XX  
 XX Query Match 100.0%; Score 2324; DB 9; Length 2324;  
 XX Best Local Similarity 100.0%; Pred. No. 0;  
 XX Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 XX 1 CGGACCGTGAATGGCCCGTAAAGATGCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
 XX 1 CGGACCGTGAATGGCCCGTAAAGATGCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
 XX 61 GCTCCGCGCAGCGTACGAGTGGCGGCGCAGCAGCAGCGGCGGCGGCGGCGGCGG 120  
 XX 61 GCTCCGCGCAGCGTACGAGTGGCGGCGCAGCAGCAGCGGCGGCGGCGGCGGCGG 120  
 XX 121 AGCATGGAGACACAGAAATGGAGGAAAGCGGACCTCTCGGGAATCTTCAATTTTCAGT 180  
 XX 121 AGCATGGAGACACAGAAATGGAGGAAAGCGGACCTCTCGGGAATCTTCAATTTTCAGT 180  
 XX 181 GGTATTGTTGATGTCATTTGCGCGTCTGGGCGTCTGGGCGTCTGGGCGTCTGGGCGT 240  
 XX 181 GGTATTGTTGATGTCATTTGCGCGTCTGGGCGTCTGGGCGTCTGGGCGTCTGGGCGT 240  
 XX 241 TTGTTGACTATAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGCTGATGCTG 300  
 XX 241 TTGTTGACTATAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGCTGATGCTG 300  
 XX 301 ATTTTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 360  
 XX 301 ATTTTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 360

Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACACTGAGCCGAGAGAGAGAGTTCCTGTGG 420  
 Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACACTGAGCCGAGAGAGAGAGTTCCTGTGG 420  
 Qy 421 AGGAGAACCCAGCATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
 Db 421 AGGAGAACCCAGCATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
 Qy 481 AAATGGTACACGCAAGACATGTTGAGGGGAGAGACTTGCACAAAGAAAGATGGACCCACAG 540  
 Db 481 AAATGGTACACGCAAGACATGTTGAGGGGAGAGACTTGCACAAAGAAAGATGGACCCACAG 540  
 Qy 541 GAGAACCAACCAAGAGGATGATGAGTTCCTTATGGCGACTGATGATGATGATGATGATG 600  
 Db 541 GAGAACCAACCAAGAGGATGATGAGTTCCTTATGGCGACTGATGATGATGATGATGATG 600  
 Qy 601 AGACCTGGAACTGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 Db 601 AGACCTGGAACTGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 Qy 661 CAGTTTCAAGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 Db 661 CAGTTTCAAGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 Qy 721 ATTCAGTGAACCAAGAGTGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 780  
 Db 721 ATTCAGTGAACCAAGAGTGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 780  
 Qy 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACTCTAGAAATGAAGGATGAGAAATCA 840  
 Db 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACTCTAGAAATGAAGGATGAGAAATCA 840  
 Qy 841 CAGAACTACTGCTCCCTGAGGATATCTCTGTAAGATTCACAGGTAATTTGTAGAG 900  
 Db 841 CAGAACTACTGCTCCCTGAGGATATCTCTGTAAGATTCACAGGTAATTTGTAGAG 900  
 Qy 901 AAGTAGCATTTTCTGTAAGAACAGCAGGAGGATGATGATGATGATGATGATGATGATG 960  
 Db 901 AAGTAGCATTTTCTGTAAGAACAGCAGGAGGATGATGATGATGATGATGATGATGATG 960  
 Qy 961 CAGATGATCCAGAACAAAGTTAAGAAAGAAAGAGCTTAACTTTTAAATTAAT 1020  
 Db 961 CAGATGATCCAGAACAAAGTTAAGAAAGAAAGAGCTTAACTTTTAAATTAAT 1020  
 Qy 1021 TTGATAAGACTATTAAGCTGAACCTGATGCTGCGAAGAACTCCGTAAGAGGGAAGAAA 1080  
 Db 1021 TTGATAAGACTATTAAGCTGAACCTGATGCTGCGAAGAACTCCGTAAGAGGGAAGAAA 1080  
 Qy 1081 TTGAGGAAGCAGTGAATGATTAAGAACTAGTAGCGAAATACCTCAGAGTCCACGAG 1140  
 Db 1081 TTGAGGAAGCAGTGAATGATTAAGAACTAGTAGCGAAATACCTCAGAGTCCACGAG 1140  
 Qy 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGATTAATGAG 1200  
 Db 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGATTAATGAG 1200  
 Qy 1201 TGCTAGCTGGAGCCATCAGACCTACCAAGAGTGGCGAGCTTACCTGATCTCCCTGAG 1260  
 Db 1201 TGCTAGCTGGAGCCATCAGACCTACCAAGAGTGGCGAGCTTACCTGATCTCCCTGAG 1260  
 Qy 1261 ACCCTGAGCTGAGTGAAGCGTCTGAGAGGCAACATTTCTAGGTCATATGA 1320  
 Db 1261 ACCCTGAGCTGAGTGAAGCGTCTGAGAGGCAACATTTCTAGGTCATATGA 1320  
 Qy 1321 GAGTTTCCCTGTTACCTCGAGAGATTAAGTCAACTATTTCCCAATGATATCTTCTTAA 1380  
 Db 1321 GAGTTTCCCTGTTACCTCGAGAGATTAAGTCAACTATTTCCCAATGATATCTTCTTAA 1380  
 Qy 1381 AAAATGACCTTTGGCGTGGGATACCTCTTGTATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
 Db 1381 AAAATGACCTTTGGCGTGGGATACCTCTTGTATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
 Qy 1441 ATGAAGAGTGTGATGTCAGACCTTAATGATGGCTTTTGTCTAAAGTCCATATGAGCTTCA 1500

1441 ATGAAGAGTCTGAGTGTGACACCTTAATGATGCTTGTCTAAGTCCATTATGGCTTCA 1500  
1501 TCCTGAAGGCACAGAAACAAATATGCTGAGAGCATCCCATATTTAAAGGAAGAAATAGAAT 1560  
1501 TCCTGAAGGCACAGAAACAAATATGCTGAGAGCATCCCATATTTAAAGGAAGAAATAGAAT 1560  
1561 CCGAGATCTCTGGCACTGATGATGGAGATTTATTTCCACCTGGGGGATGCCATCAGCA 1620  
1561 CCGAGATCTCTGGCACTGATGATGGAGATTTATTTCCACCTGGGGGATGCCATCAGCA 1620  
1621 GGGTTGGGAACAAAGAGCATATAGTGTATAGCTTGGGCAACAGAGAGGACACTTTG 1680  
1621 GGGTTGGGAACAAAGAGCATATAGTGTATAGCTTGGGCAACAGAGAGGACACTTTG 1680  
1681 CATCTGTCTGGCAACCGCTCTACATGATGATGATGATGATGATGATGATGATGATGATG 1740  
1681 CATCTGTCTGGCAACCGCTCTACATGATGATGATGATGATGATGATGATGATGATGATG 1740  
1741 CCCCAAGAAACCGGCTTACACAGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
1741 CCCCAAGAAACCGGCTTACACAGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
1801 TCCGAGATGAAGGCTTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
1801 TCCGAGATGAAGGCTTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
1861 AAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920  
1861 AAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920  
1921 ATGAAATGCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980  
1921 ATGAAATGCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980  
1981 CAACAGGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040  
1981 CAACAGGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040  
2041 GCGCGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
2041 GCGCGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
2101 AGGAGGCTGCAAGATTCATGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160  
2101 AGGAGGCTGCAAGATTCATGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160  
2161 TCATCTTTGATGATCTCTTTGAGCAGGAGGATGAGCAGGATGAGCAGGATGAGCAGGATG 2220  
2161 TCATCTTTGATGATCTCTTTGAGCAGGAGGATGAGCAGGATGAGCAGGATGAGCAGGATG 2220  
2221 TATTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280  
2221 TATTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280  
2281 CAAATTTAGCATGAATTCATGCAAGTGTGGAACTCTGGAGAGA 2324  
2281 CAAATTTAGCATGAATTCATGCAAGTGTGGAACTCTGGAGAGA 2324

RESULT 4  
ADP23914  
ID ADP23914 standard; cdna; 2452 BP.  
XX  
AC ADP23914;  
AC ADP23914;  
DT 18-NOV-2004 (first entry)  
XX  
DE PRO polypeptide encoding cdna SEQ ID NO:1092.  
XX  
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
KW antipsoaric; antiallergic; antiasthmatic; hepatotropic; respiratory;

gene therapy; immune system.  
Unidentified.  
WO2004041170-A2.  
XX  
XX 21-MAY-2004.  
XX  
XX 30-OCT-2003; 2003WO-US034312.  
XX  
XX 01-NOV-2002; 2002US-0423394P.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
XX Wu TD;  
XX  
XX WPI: 2004-419628/39.  
XX P-PSDB; ADF23915.  
XX  
XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
XX erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
XX renal disease, or demyelinating diseases of the central or peripheral  
XX nervous system.  
XX  
XX Claim 1; SEQ ID NO 1092; 2940pp; English.  
XX  
XX The invention relates to a novel isolated nucleic acid and the PRO  
XX polypeptide encoded by it. A protein of the invention has  
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
XX osteopathic, antidiabetic, dermatological, antipsoaric, antiallergic,  
XX antisthmatic, hepatotropic, and respiratory activity. A polynucleotide  
XX of the invention may have a use in gene therapy. The PRO polypeptide, its  
XX agonist, antagonist, or antibody that specifically binds to the  
XX polypeptide is useful for treating an immune related disorder such as  
XX systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,  
XX juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
XX idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune  
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, immune-mediated renal  
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
XX disease, a demyelinating disease of the central or peripheral nervous  
XX system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
XX a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
XX disease, infectious or autoimmune chronic active hepatitis, primary  
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
XX disease, an autoimmune or immune-mediated skin disease, a bullous skin  
XX disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
XX disease, asthma, allergic rhinitis, atopic dermatitis, food  
XX hypersensitivity, urticaria, an immunologic disease of the lung,  
XX eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
XX pneumonitis, a transplantation associated disease, graft rejection or  
XX graft-versus-host disease. The present sequence encodes a PRO protein of  
XX the invention.

Sequence 2452 BP; 779 A; 507 C; 657 G; 509 T; 0 U; 0 Other;  
Query Match 99.9%; Score 2320.8; DB 13; Length 2452;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGACCGTGTCAATGCGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 67 CGGACCGTGTCAATGCGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 126  
QY 61 GCTCGGAGCGGTAGCAGCAGTGGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 120  
DB 127 GCTCGGAGCGGTAGCAGCAGTGGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 186  
QY 121 AGCATGAGGAGCACAAAGAAATGGGAGGAAGCGGACCTCTCGGGAACCTTCATCTTCAGGT 180  
DB 187 AGCATGAGGAGCACAAAGAAATGGGAGGAAGCGGACCTCTCGGGAACCTTCATCTTCAGGT 246

Qy 181 GGTATATGTTGATTCGATTTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
Db 247 GGTATATGTTGATTCGATTTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 306  
Qy 241 TTGTTCACTATGAGGAAGTCTTAGGAAAACTAGGAATCTATGATCTCATGGTGTATGGAG 300  
Db 307 TTGTTGACTATGAGGAAGTCTTAGGAAAACTAGGAATCTATGATCTCATGGTGTATGGAG 366  
Qy 301 ATTTTGATGTGATGATGCCAAAGTTTTTATTAGGACTTAAGAGAGATCTACTTCAGAGC 360  
Db 367 ATTTTGATGTGATGATGCCAAAGTTTTTATTAGGACTTAAGAGAGATCTACTTCAGAGC 426  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTTGAGCCACACTGAGCCGAGAGCAGGTTCTCTGTGG 420  
Db 427 CAGCAGTCCCGCCAGAGAGGCTTGAGCCACACTGAGCCGAGAGCAGGTTCTCTGTGG 486  
Qy 421 AGGCAGAACCCGAGATATCGAAGATGAAGCAAAAGAACTTCACTCCCTTCTCCATG 480  
Db 487 AGGCAGAACCCGAGATATCGAAGATGAAGCAAAAGAACTTCACTCCCTTCTCCATG 546  
Qy 481 AAATGTTACGCGAAGACATGTTGAGGAGAAAGACTTGCAACAAAGAGATGAGCCACAG 540  
Db 547 AAATGTTACGCGAAGACATGTTGAGGAGAAAGACTTGCAACAAAGAGATGAGCCACAG 606  
Qy 541 GAGAACCAACAAAGAGAGATGATGAGTTTCTTATGGCACTGATGATGATGATGATG 600  
Db 607 GAGAACCAACAAAGAGAGATGATGAGTTTCTTATGGCACTGATGATGATGATGATG 666  
Qy 601 AGACCTCGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAGAGAG 660  
Db 667 AGACCTCGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAGAGAG 726  
Qy 661 CAGTTTCAACAGCTGAATCAGGATATGGAAGAGATGATGCTGAGCAGAGAAATCCAG 720  
Db 727 CAGTTTCAACAGCTGAATCAGGATATGGAAGAGATGATGCTGAGCAGAGAAATCCAG 786  
Qy 721 ATTCAGTGAACAGTAGTAGAAGATGAAGATTCACCATGATACAGATGATGATGAT 780  
Db 787 ATTCAGTGAACAGTAGTAGAAGATGAAGATTCACCATGATACAGATGATGATGAT 846  
Qy 781 ACCAAGTCTATGAGGACAGCAGTATGATGATGATGATGATGATGATGATGATGAT 840  
Db 847 ACCAAGTCTATGAGGACAGCAGTATGATGATGATGATGATGATGATGATGATGAT 906  
Qy 841 CAGAGTAACTGCTCCCTGAGGATAATCTCTGTAGAGATTCACAGGTAATTTGTAAG 900  
Db 907 CAGAGTAACTGCTCCCTGAGGATAATCTCTGTAGAGATTCACAGGTAATTTGTAAG 966  
Qy 901 AAGTAAGCATTTTCTGTGGAGAACAGCAGGAGATACCAACAGAGAAACAAATAGAAAA 960  
Db 967 AAGTAAGCATTTTCTGTGGAGAACAGCAGGAGATACCAACAGAGAAACAAATAGAAAA 1026  
Qy 961 CAGATGATCAGAGAAACAAAGTTTAGAAGAGAGCCTTAACCTTTTAATAAT 1020  
Db 1027 CAGATGATCAGAGAAACAAAGTTTAGAAGAGAGCCTTAACCTTTTAATAAT 1086  
Qy 1021 TTGATAGACTATTAAAGCTGAACCTTGATGCTGAGAGAAACCTCCGTAAAAAGGAAAA 1080  
Db 1087 TTGATAGACTATTAAAGCTGAACCTTGATGCTGAGAGAAACCTCCGTAAAAAGGAAAA 1146  
Qy 1081 TTGAGGAGCAGTGAATGCAATTAAGAACTAGTAGCAAAATACCTCAGAGTCCAGAG 1140  
Db 1147 TTGAGGAGCAGTGAATGCAATTAAGAACTAGTAGCAAAATACCTCAGAGTCCAGAG 1206  
Qy 1141 CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGATGATGAGG 1200  
Db 1207 CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGATGATGAGG 1266  
Qy 1201 TGCTAGTGAGCCATCCGAGACCTACCAAGAGTGGCCAGCCTACTGATGCTCCCTGAG 1260  
Db 1267 TGCTAGTGAGCCATCCGAGACCTACCAAGAGTGGCCAGCCTACTGATGCTCCCTGAG 1326  
Qy 1261 ACTGCTGAAGCTGATTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320

Db 1327 ACTGCTGAAGCTGATTTGAAGCGTCTGAGCAGGCAACAATTTCTAGGTCAATGA 1386  
Qy 1321 GAGGTTCCCTGCTTACCTCGACAGATTAAGTTCAACTATTTCCCAATGATCTTCTTAA 1380  
Db 1387 GAGGTTCCCTGCTTACCTCGACAGATTAAGTTCAACTATTTCCCAATGATCTTCTTAA 1446  
Qy 1381 AAAATGACCTTTGGCGTGGGATACCTCTGATAGGAGATAATGCAATGCAAGAAAGTTT 1440  
Db 1447 AAAATGACCTTTGGCGTGGGATACCTCTGATAGGAGATAATGCAATGCAAGAAAGTTT 1506  
Qy 1441 ATGAAGAGGCTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
Db 1507 ATGAAGAGGCTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1566  
Qy 1501 TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAAT 1560  
Db 1567 TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAAT 1626  
Qy 1561 CGGAGATCTCTGGCACTGATGATGGGAGATTTTATTTTCCACCTGGGGGATGCCATGSCAG 1620  
Db 1627 CGGAGATCTCTGGCACTGATGATGGGAGATTTTATTTTCCACCTGGGGGATGCCATGSCAG 1686  
Qy 1621 GGGTTGGGAAACAAAGAGGCAATATAAGTGGTATGAGCTTTGGGCAACAGAGAGGACATTTG 1680  
Db 1687 GGGTTGGGAAACAAAGAGGCAATATAAGTGGTATGAGCTTTGGGCAACAGAGAGGACATTTG 1746  
Qy 1681 CATCTGTCTGGCAACGCTCACTCTACATGTGAATGCACTGAGCAACAGCCTTGGTGA 1740  
Db 1747 CATCTGTCTGGCAACGCTCACTCTACATGTGAATGCACTGAGCAACAGCCTTGGTGA 1806  
Qy 1741 CCCCAGAGAACCGGCTACACAGAGTTGTAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
Db 1807 CCCCAGAGAACCGGCTACACAGAGTTGTAAGTCTTTAGAAAAGAACTGGAAGTTAA 1866  
Qy 1801 TCCGAGATGAAGGCTTGCAGTGATGAATGAAGCAAAAGCTCTCTCTCTGCTGAGGATG 1860  
Db 1867 TCCGAGATGAAGGCTTGCAGTGATGAATGAAGCAAAAGCTCTCTCTCTGCTGAGGATG 1926  
Qy 1861 AAAACCTGAGGAAAAAGGGAAGGCACTGGAGCCAGTTTACGCTGTGGGCAACAGAGAGAA 1920  
Db 1927 AAAACCTGAGGAAAAAGGGAAGGCACTGGAGCCAGTTTACGCTGTGGGCAACAGAGAGAA 1986  
Qy 1921 ATGAAAATGCTGCAAGAGGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAGA 1980  
Db 1987 ATGAAAATGCTGCAAGAGGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAGA 2046  
Qy 1981 CAACAGATGCAAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGGAACCTCACGTGT 2040  
Db 2047 CAACAGATGCAAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGGAACCTCACGTGT 2106  
Qy 2041 GGCCGCAACACAGGCCCCAACCTGCAAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
Db 2107 GGCCGCAACACAGGCCCCAACCTGCAAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2166  
Qy 2101 AGGAAGGCTGCAAGATTCGATGTGCCAAAGAGACAGGACCTGGGAGAGAGAGGAGTGC 2160  
Db 2167 AGGAAGGCTGCAAGATTCGATGTGCCAAAGAGACAGGACCTGGGAGAGAGAGGAGTGC 2226  
Qy 2161 TCATCTTTGATGATCTCTTTTGGAGCAGAGTATGAGGATGCTCATCTTTTCGGGCTGA 2220  
Db 2227 TCATCTTTGATGATCTCTTTTGGAGCAGAGTATGAGGATGCTCATCTTTTCGGGCTGA 2286  
Qy 2221 TATTCATCTGATGATGTGTGCATCCGGAACCTGACACCAACAGAGAGAGAGGAGGCTTCCAG 2280  
Db 2287 TATTCATCTGATGATGTGTGCATCCGGAACCTGACACCAACAGAGAGAGAGGAGGCTTCCAG 2346  
Qy 2281 CAATTTAGCATGAATTCATCAAGCTTTGGAAACTCTGAGAGA 2324  
Db 2347 CAATTTAGCATGAATTCATCAAGCTTTGGAAACTCTGAGAGA 2390

187	AGCATGGAGGACACAAGAATGGGAGGAAAGCGGAGCTCTCAGGAACTTCATTCTTCCAGT	246
181	GGTTTATGGTCAATTCATTTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC	240
247	GGTTTATGGTCAATTCATTTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC	306
241	TTGTTGACTATGAGGAAAGTTCTAGGAAAACTAGAAATCTATGATGCTGTGATGGATGGAG	300
307	TTGTTGACTATGAGGAAAGTTCTAGGAAAACTAGAAATCTATGATGCTGTGATGGATGGAG	366
301	ATTTTATGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	360
367	ATTTTATGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	426
361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGCTTCTGTGG	420
427	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGCTTCTGTGG	486
421	AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	480
487	AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	546
481	AAATGTTACACGACAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	540
547	AAATGTTACACGACAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	606
541	GAGAACCAACAAGAGGATGATGATTTCTTATGCGGACATGATGATGATGATGATGATGATGAT	600
607	GAGAACCAACAAGAGGATGATGATTTCTTATGCGGACATGATGATGATGATGATGATGATGAT	666
601	AGACCTGGAACCTGAAATATCTCATGAGAAACCGAGCATAGTTACCACTGGGAGAGA	660
667	AGACCTGGAACCTGAAATATCTCATGAGAAACCGAGCATAGTTACCACTGGGAGAGA	726
661	CAGTTTCAACAAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG	720
727	CAGTTTCAACAAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG	786
721	ATTCCAGTGAACAGTAGTAGAAGATGAAGATTCACCATGATATACAGATGATGATGATGAT	780
787	ATTCCAGTGAACAGTAGTAGAAGATGAAGATTCACCATGATATACAGATGATGATGATGAT	846
781	ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA	840
847	ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA	906
841	CAGAAGTAACTGCTCCCTGAGGATATCTCTGAGAGATTCCTGAGGATTCACAGGTAATTTGAGAAG	900
907	CAGAAGTAACTGCTCCCTGAGGATATCTCTGAGAGATTCCTGAGGATTCACAGGTAATTTGAGAAG	966
901	AAGTAACTGCTCCCTGAGGATATCTCTGAGAGATTCCTGAGGATTCACAGGTAATTTGAGAAG	960
967	AAGTAACTGCTCCCTGAGGATATCTCTGAGAGATTCCTGAGGATTCACAGGTAATTTGAGAAG	1026
961	CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATTAAT	1020
1027	CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATTAAT	1086
1021	TTGATAAGACTATTTAAAGCTGAACTTGTATGCTGAGAAACCTCCGTAAGAGGAGAAAA	1080
1087	TTGATAAGACTATTTAAAGCTGAACTTGTATGCTGAGAAACCTCCGTAAGAGGAGAAAA	1146
1081	TTGAGGAGAGCTGAACTGAACTTTAAAGAACTAGTACCAAAATACCTCCAGAGTCCAGAG	1140
1147	TTGAGGAGAGCTGAACTGAACTTTAAAGAACTAGTACCAAAATACCTCCAGAGTCCAGAG	1206
1141	CAAGATATGGGAAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGATTAATGAGG	1200
1207	CAAGATATGGGAAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGATTAATGAGG	1266
1201	TGCTACGTTGGAGCCATCCAGACCTTACCAAGAGGTTGGCCAGGCTACCTGATGCTCCCTGAG	1260
1267	TGCTACGTTGGAGCCATCCAGACCTTACCAAGAGGTTGGCCAGGCTACCTGATGCTCCCTGAG	1326

ADR97347	Db
ID ADR97347 standard; DNA; 2452 BP.	Qy
AC ADR97347;	Db
XX 02-DEC-2004 (first entry)	Qy
XX Human ASPH DNA, an apoptosis related target Seq 55.	Db
XX gene; ds; human; apoptosis; cancer; inflammation; autoimmune;	Qy
XX neurodegenerative disorder; cytostatic; antiinflammatory;	Db
XX immunosuppressive; neuroprotective; gene therapy; ASPH;	Qy
XX aspartate beta hydroxylase.	Db
XX Homo sapiens.	Qy
XX WO2004078783-A2.	Db
XX 16-SEP-2004.	Qy
XX 05-MAR-2004; 2004WO-GB000957.	Db
XX 07-MAR-2003; 2003GB-00005267.	Qy
XX (BIRX-) EIRX THERAPEUTICS LTD.	Db
XX Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;	Qy
XX WPI; 2004-662402/64.	Db
XX P-PSDB; ADR97348.	Qy
XX Identifying an agent that modulates the function of an apoptosis-	Db
XX associated polypeptide, useful for diagnosing or treating e.g. cancer,	Qy
XX comprises comparing the binding of the polypeptide to the candidate agent	Db
XX and to a control agent.	Qy
XX Claim 2; SEQ ID NO 55; 304pp; English.	Db
XX This invention relates to novel agents that modulates the function of	Qy
XX human apoptosis-associated proteins specified within the specification.	Db
XX Specifically, it refers to a method for the identification of target	Qy
XX genes whose expression is correlated with an early stage in the	Db
XX regulation of apoptosis. The present invention describes a method of	Qy
XX contacting either candidate agents or control agents to the target genes	Db
XX and assessing the difference of binding and inhibitory activity, where	Qy
XX the candidate agent is selected from a low molecular weight organic	Db
XX molecule, an antibody or fragment thereof, an antisense oligonucleotide,	Qy
XX a small inhibitory dRNA, or a ribozyme. As such, the compositions and	Db
XX methods are useful for diagnosing and treating diseases or conditions	Qy
XX associated with abnormal apoptosis in mammalian tissues, such as cancer,	Db
XX inflammation, autoimmune or neurodegenerative disorders. Accordingly,	Qy
XX they exhibit cytostatic, antiinflammatory, immunosuppressive and	Db
XX neuroprotective activities. These may also be used for drug screening	Qy
XX purposes and in gene therapy. This polynucleotide sequence is a human	Db
XX target gene associated with the regulation of apoptosis that can be	Qy
XX modulated by novel agents of the invention.	Db
XX Sequence 2452 BP; 779 A; 507 C; 657 G; 509 T; 0 U; 0 Other;	Qy
XX Query Match 99.9%; Score 2320.8; DB 13; Length 2452;	Db
XX Best Local Similarity 99.9%; Pred. No. 0;	Qy
XX Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Db
Qy 1 CGGACCTGCAATGCGCCACGCTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG 60	
Db 67 CGGACCTGCAATGCGCCACGCTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG 126	
Qy 61 GCTCCGCGCAGCGGTAGCAGCAGGTCCGGCAGCAGCAGCCCGGGCCCGGAGAGACAA 120	
Db 127 GCTCCGCGCAGCGGTAGCAGCAGGTCCGGCAGCAGCAGCCCGGGCCCGGAGAGACAA 186	
Qy 121 AGCATGAGGACACAGAATGGGAGGAAAGCGGAGCTCTCGGAACTTCATTCTTCCAGT 180	

QY 1261 ACTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
Db |||||  
QY 1327 ACTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCATATGA 1386  
Db |||||  
QY 1321 GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCACACTATTTCCCAATGATACCTCTTAA 1380  
Db |||||  
QY 1387 GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCACACTATTTCCCAATGATACCTCTTAA 1446  
Db |||||  
QY 1381 AAAATGACCTTGGCGTGGATACCTCTGATAGGAGATTAATGACATGCAAGCAAGATTTT 1440  
Db |||||  
QY 1447 AAAATGACCTTGGCGTGGATACCTCTGATAGGAGATTAATGACATGCAAGCAAGATTTT 1506  
Db |||||  
QY 1441 ATCAAGAGTCTGAGTGTGACACCTTAATGATGGCTTTTCTAAAGTCAATATGCTTCA 1500  
Db |||||  
QY 1507 ATGAAGAGTCTGAGTGTGACACCTTAATGATGGCTTTTCTAAAGTCAATATGCTTCA 1566  
Db |||||  
QY 1501 TCCTGAAGGCACAGAACAAATTTCTGAGAGCATCCCATATTTAAAGGAAGGATAGAT 1560  
Db |||||  
QY 1567 TCCTGAAGGCACAGAACAAATTTCTGAGAGCATCCCATATTTAAAGGAAGGATAGAT 1626  
Db |||||  
QY 1561 CCGAGATCTCGGCACCTGATGATGGAGATTTTATTTCCACCTGGGGATGCCATCAGA 1620  
Db |||||  
QY 1627 CCGAGATCTCGGCACCTGATGATGGAGATTTTATTTCCACCTGGGGATGCCATCAGA 1686  
Db |||||  
QY 1621 GGGTTGGGACAAAGAGGCATATAGTGTATGAGCTTTGGGCACAGAGAGCACTTTG 1680  
Db |||||  
QY 1687 GGGTTGGGACAAAGAGGCATATAGTGTATGAGCTTTGGGCACAGAGAGCACTTTG 1746  
Db |||||  
QY 1681 CATCTGTCTGGCAACCGCTCACTCAATGTGAATGGAAGTGAAGCAGACAGCTTTGTGA 1740  
Db |||||  
QY 1747 CATCTGTCTGGCAACCGCTCACTCAATGTGAATGGAAGTGAAGCAGACAGCTTTGTGA 1806  
Db |||||  
QY 1741 CCCCAAGAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db |||||  
QY 1807 CCCCAAGAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1866  
Db |||||  
QY 1801 TCCAGATGAGGCGCTTCAGTGTGATGATGAAGCAAGTCTTCTTCCCTGCTGAGGATG 1860  
Db |||||  
QY 1867 TCCAGATGAGGCGCTTCAGTGTGATGATGAAGCAAGTCTTCTTCCCTGCTGAGGATG 1926  
Db |||||  
QY 1861 AAAACCTGAGGAAAGAGGAGCTGAGCCAGTTTCACTGTCTGGCAGCAAGAGAGAA 1920  
Db |||||  
QY 1927 AAAACCTGAGGAAAGAGGAGCTGAGCCAGTTTCACTGTCTGGCAGCAAGAGAGAA 1986  
Db |||||  
QY 1921 ATGAAATGCTGCAAGAGAGCTCCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAGA 1980  
Db |||||  
QY 1987 ATGAAATGCTGCAAGAGAGCTCCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAGA 2046  
Db |||||  
QY 1981 CAACAGGATGCAAGAGAGAGATCAAAATTTCCATGATGCAACCCCGGACTCAGGTG 2040  
Db |||||  
QY 2047 CAACAGGATGCAAGAGAGAGATCAAAATTTCCATGATGCAACCCCGGACTCAGGTG 2106  
Db |||||  
QY 2041 GGGCGCACACGGGCGCCAAACTGCGAGCTTCGATGCACTTGGGCTTGTGATTTCCCA 2100  
Db |||||  
QY 2107 GGGCGCACACGGGCGCCAAACTGCGAGCTTCGATGCACTTGGGCTTGTGATTTCCCA 2166  
Db |||||  
QY 2101 AGGAGGCTGCAAGATTCGATGTGCCAAACGAGACCAAGACCTGGGAGGAAGCAAGTGC 2160  
Db |||||  
QY 2167 AGGAGGCTGCAAGATTCGATGTGCCAAACGAGACCAAGACCTGGGAGGAAGCAAGTGC 2226  
Db |||||  
QY 2161 TCATCTTTGATGACTCTTTTTCAGCACGAGGTATGGCAGGATGCCTCATCTTTTCCGGCTGA 2220  
Db |||||  
QY 2227 TCATCTTTGATGACTCTTTTTCAGCACGAGGTATGGCAGGATGCCTCATCTTTTCCGGCTGA 2286  
Db |||||  
QY 2221 TATTATCTGTGATGTGGCATCCGAACTGACACACACAGAGAGAGAGCCCTTCCAG 2280  
Db |||||  
QY 2287 TATTATCTGTGATGTGGCATCCGAACTGACACACACAGAGAGAGAGCCCTTCCAG 2346  
Db |||||  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGTGAGAGA 2324  
Db |||||  
QY 2347 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGTGAGAGA 2390  
Db |||||

RESULT 6  
ADJ56257 standard; cDNA; 5358 BP.  
XX AC ADJ56257;  
XX DT 06-MAY-2004 (first entry)  
XX XX Bovine cDNA differentially expressed in MYCN activated cells SeqID 63.  
XX DE bovine; differential expression; transactivator; proto-oncogene;  
XX KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; se;  
XX KW MYCN activated cell.  
XX OS Bos taurus.  
XX XX US2003119009-A1.  
XX PN 26-JUN-2003.  
XX PD 25-FEB-2002; 2002US-00084817.  
XX PF 23-FEB-2001; 2001US-0270784P.  
XX PR (STUA/) STUART S G.  
XX PA (NUCH/) NUCHTERN J G.  
XX PA (PLON/) PLON S E.  
XX PA (SHOH/) SHOHET J M.  
XX PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;  
XX XX WPI; 2003-635698/60.  
XX XX New genes regulated by MYCN activation, useful in gene therapy,  
XX PT particularly for treating a subject with e.g. neuroblastoma or other  
XX PT cancers, or for diagnosing, staging or monitoring the treatment of the  
XX PT cancer.  
XX PS Claim 1; SEQ ID NO 63; 27pp; English.  
XX XX This invention relates to novel isolated cDNAs that are differentially  
XX CC expressed in MYCN activated cells. Specifically, it refers to  
XX CC polynucleotide sequences that exhibit differential expression patterns in  
XX CC cells activated by the transactivator MYCN, where MYCN is a proto-  
XX CC oncogene that is amplified in neuroblastoma cells and is common in small  
XX CC cell lung cancers. The present invention describes these cDNA molecules  
XX CC as useful for in hybridisation assays to detect expression of nucleic  
XX CC acids (or complementary nucleic acids) in a present in a given sample, as  
XX CC well as for screening assays by identifying molecules or compounds that  
XX CC specifically bind the cDNA as a ligand and modulate function or activity.  
XX CC Accordingly, these compositions exhibit cytostatic activity and can also  
XX CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA  
XX CC that is differentially expressed in MYCN activated cells and can also  
XX CC exemplification of the invention. NOTE: This sequence does not appear in  
XX CC the printed specification but has been obtained in electronic format from  
XX CC the US Patent Office at  
XX CC ftp.sequata.uspto.gov/sequence.html?DocID=20030119009.  
XX SQ Sequence 5358 BP; 1684 A; 1009 C; 1204 G; 1461 T; 0 U; 0 Other;

Query Match 99.8%; Score 2319.2; DB 10; Length 5358;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATCGCCACGTAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 60  
Db |||||  
QY 258 CGGACCGTGCATCGCCACGTAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 317  
Db |||||  
QY 61 GCTCCGGCAGCGGTAGCAGCGTGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 120  
Db |||||  
QY 318 GCTCCGGCAGCGGTAGCAGCGTGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 377  
Db |||||

QY	121	AGCATGGAGGACACAAAGATGGGAGGAAAGGCGGACTCTCGGGAACCTTCATTCTTCCAGT	180
Db	378	AGCATGGAGGACACAAAGATGGGAGGAAAGGCGGACTCTCAGGAACTTCATTCTTCCAGT	437
QY	181	GGTTATGGTGATGTCATCTCTGGCGTCTGGACATCTGTAGCTGTCGTTTGGTTGATC	240
Db	438	GGTTATGGTGATGTCATCTCTGGCGTCTGGACATCTGTAGCTGTCGTTTGGTTGATC	497
QY	241	TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGTTGATGGAG	300
Db	498	TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGTTGATGGAG	557
QY	301	ATTTTGTGATGATGATGTCCTCAAGTTTATAGGAATTTAAAGAGAGATCTACTTCAAGC	360
Db	558	ATTTTGTGATGATGATGTCCTCAAGTTTATAGGAATTTAAAGAGAGATCTACTTCAAGC	617
QY	361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG	420
Db	618	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG	677
QY	421	AGGAGAACCCAGAAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCTCCATG	480
Db	678	AGGAGAACCCAGAAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCTCCATG	737
QY	481	AAATGGTACACGAGAACTGTTGAGGAGAGACTTGCACAGAGAGATGGACCCACAG	540
Db	738	AAATGGTACACGAGAACTGTTGAGGAGAGACTTGCACAGAGAGATGGACCCACAG	797
QY	541	GAGAACCAACACAGAGATGATGATGTTCTTATGGGACTGATGATGATGATGATGATG	600
Db	798	GAGAACCAACACAGAGATGATGATGTTCTTATGGGACTGATGATGATGATGATGATG	857
QY	601	AGACCTTGGAACTTCAATGAAGAAACCGAGCATAGTATCCACGTCGAAGAGA	660
Db	858	AGACCTTGGAACTTCAATGAAGAAACCGAGCATAGTATCCACGTCGAAGAGA	917
QY	661	CAGTTTCAACAGACTGTAATCAGGATGGAAGAGATGATCTGAGCAGGAAATCCAG	720
Db	918	CAGTTTCAACAGACTGTAATCAGGATGGAAGAGATGATCTGAGCAGGAAATCCAG	977
QY	721	ATTCCAGTGAACCACTAGTAGAAGTGAAGATGCAACATGATACATGATGATGAT	780
Db	978	ATTCCAGTGAACCACTAGTAGAAGTGAAGATGCAACATGATACATGATGATGAT	1037
QY	781	ACCAAGTCTATGAGGAACAGCAGATGATGATGATGATGATGATGATGATGATGAT	840
Db	1038	ACCAAGTCTATGAGGAACAGCAGATGATGATGATGATGATGATGATGATGATGAT	1097
QY	841	CAGAGTAACTGCTCCCTGAGGATTAATCTGTAGAGATTAATCTGTAGAGATTAAT	900
Db	1098	CAGAGTAACTGCTCCCTGAGGATTAATCTGTAGAGATTAATCTGTAGAGATTAAT	1157
QY	901	AGTAAAGCAATTTTCTGTGGAAGAACAGCAGAGATGATGATGATGATGATGATGAT	960
Db	1158	AGTAAAGCAATTTTCTGTGGAAGAACAGCAGAGATGATGATGATGATGATGATGAT	1217
QY	961	CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAGAAAGCTTAACTTTTAAATAAT	1020
Db	1218	CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAGAAAGCTTAACTTTTAAATAAT	1277
QY	1021	TTGATAAGCACTTAAAGCTGAACTTTGATGCTGACAGAAAACTCCGTAAAGGGGAAAA	1080
Db	1278	TTGATAAGCACTTAAAGCTGAACTTTGATGCTGACAGAAAACTCCGTAAAGGGGAAAA	1337
QY	1081	TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATCCCTCAGAGTCCACGAG	1140
Db	1338	TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATCCCTCAGAGTCCACGAG	1397
QY	1141	CAGATATGGAAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGATTAATGAGG	1200
Db	1398	CAGATATGGAAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGATTAATGAGG	1457
QY	1201	TGCTACGTGGAGCAATCGAGACCTTACCAAGAGTGGCCAGCCTACCTGATGCTCCCTGCAG	1260

Db	1458	TGCTACGTGGAGCAATCGAGACCTTACCAAGAGTGGCCAGCCTACCTGATGCTCCCTGCAG	1517
QY	1261	ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGCAAAATTTCTAGGTCAATGA	1320
Db	1518	ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGCAAAATTTCTAGGTCAATGA	1577
QY	1321	GAGGTTCCCTCTTACCTGCGAGAGATTTAGTTCAACTATTTTCCCAATGATATTCCTTAA	1380
Db	1578	GAGGTTCCCTCTTACCTGCGAGAGATTTAGTTCAACTATTTTCCCAATGATATTCCTTAA	1637
QY	1381	AAAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATTAATGAACAATGCAAGAAAGTTT	1440
Db	1638	AAAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATTAATGAACAATGCAAGAAAGTTT	1697
QY	1441	ATGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1500
Db	1698	ATGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1757
QY	1501	TCCTGAAGGCAAGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1560
Db	1758	TCCTGAAGGCAAGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1817
QY	1561	CCGAGATCTCTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1620
Db	1818	CCGAGATCTCTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1877
QY	1621	GGGTTGGGAAACAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1680
Db	1878	GGGTTGGGAAACAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1937
QY	1681	CATCTGTCTGGCAAGCTCTCTCAATGATGATGATGATGATGATGATGATGATGATGAT	1740
Db	1938	CATCTGTCTGGCAAGCTCTCTCAATGATGATGATGATGATGATGATGATGATGATGAT	1997
QY	1741	CCCCAAAGAAAGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
Db	1998	CCCCAAAGAAAGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	2057
QY	1801	TCCGAGATGAAGGCTTGCAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA	1860
Db	2058	TCCGAGATGAAGGCTTGCAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA	2117
QY	1861	AAAACTGAGGAAAAAGGGCTGAGCCAGTTTCAAGTGTGGCAGCAAGGAAGAA	1920
Db	2118	AAAACTGAGGAAAAAGGGCTGAGCCAGTTTCAAGTGTGGCAGCAAGGAAGAA	2177
QY	1921	ATGAAATGCTCTCAAGAGGCTCTTAAACCTGTACTTACTAGAAAAAGTTTCCCGAGA	1980
Db	2178	ATGAAATGCTCTCAAGAGGCTCTTAAACCTGTACTTACTAGAAAAAGTTTCCCGAGA	2237
QY	1981	CAACAGGATCGAAGAGGACAGATCAATATTTCCATCATGCAACCCCGGACCTCAGGTG	2040
Db	2238	CAACAGGATCGAAGAGGACAGATCAATATTTCCATCATGCAACCCCGGACCTCAGGTG	2297
QY	2041	GGCGGACACAGGCCCCACAACTGCAAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA	2100
Db	2298	GGCGGACACAGGCCCCACAACTGCAAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA	2357
QY	2101	AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACCGAGACCTGGGAGGAAGGAGTGC	2160
Db	2358	AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACCGAGACCTGGGAGGAAGGAGTGC	2417
QY	2161	TCATCTTTTGTGATGATCTCTTTTGTAGCAGGATGATGAGGAGGATGCTCATCTTCCGGCTGA	2220
Db	2418	TCATCTTTTGTGATGATCTCTTTTGTAGCAGGATGATGAGGAGGATGCTCATCTTCCGGCTGA	2477
QY	2221	TATTATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2280
Db	2478	TATTATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2537
QY	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTCTGGAGAGA	2324



Db 2538 CAATTAGCATGATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2581

RESULT 7

ABX04178

ID ABX04178 standard; cDNA; 2680 BP.

XX

AC ABX04178;

XX

DT 10-JAN-2003 (first entry)

XX

DE Human mRNA differentially expressed in mesenchymal cells #25.

XX

KW Human; ss; gene; skeletal growth; cartilage degeneration disorder;

XX

KW chondroblastic phenotype; mesenchymal cell; cartilage formation;

XX

KW bone formation; arthritis; osteoarthritis; rheumatoid arthritis;

XX

KW gout arthritis; adjuvant arthritis; arthritis deformans; antipath;

XX

KW infectious arthritis; osteochondrosis; RDA; antiarthritic; osteopathic;

XX

KW antirheumatic; antiinflammatory; representational difference analysis.

XX

OS Homo sapiens.

XX

PN WO200271927-A2.

XX

PD 19-SEP-2002.

XX

PF 12-MAR-2002; 2002WO-US007787.

XX

PR 12-MAR-2001; 2001US-0274980P.

XX

PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.

XX

PI Yates KE, Mizuno S, Glowacki J;

XX

DR WPI; 2002-723276/78.

XX

PT New nucleic acid molecules capable of promoting chondrogenesis, useful

PT for diagnosing and treating cartilaginous tissue degeneration conditions,

PT e.g. osteoarthritis, rheumatoid arthritis, gout arthritis, or

PT osteochondrosis.

XX

PS Claim 33; Page 123; 153pp; English.

XX

CC The invention relates to new isolated nucleic acid molecule comprising a

CC nucleic acid molecule consisting of a gene differentially expressed in

CC cells undergoing differentiation from mesenchymal cell to a

CC chondroblastic phenotype, or hybridising under stringent conditions to

CC them (or their fragments). Also included are expression vectors,

CC transformed host cells, expressed polypeptides or peptide fragments

CC (which induce differentiation of a mesenchymal cell and may be used as an

CC immunogen), binding partners of the polypeptides, a method for

CC identifying an agent useful in modulating mesenchymal cell

CC differentiation induction activity of a molecule, a method of diagnosing

CC a condition characterized by aberrant expression of a nucleic acid

CC molecule or its expression product; a method for determining regression,

CC progression or onset of cartilaginous tissue degeneration condition in a

CC subject characterized by aberrant expression of a nucleic acid molecule

CC or its expression product, a method for treating a cartilaginous tissue

CC degeneration condition, a method for treating a subject to reduce the

CC risk of cartilaginous tissue degeneration condition developing in the

CC subject, a method for identifying a candidate agent for treating in the

CC cartilaginous tissue degeneration condition, and a solid-phase nucleic

CC acid molecule array consisting essentially of a set of nucleic acid

CC molecule as cited above (or known from known genes shown to be

CC differentially expressed in developing mesenchymal cells using the

CC technique of representational difference analysis, RDA), its expression

CC products or fragments, fixed to a solid substrate, the nucleic acids,

CC polypeptides and agents are useful for treating cartilaginous tissue

CC degeneration conditions such as osteoarthritis, rheumatoid arthritis,

CC gout arthritis, adjuvant arthritis, arthritis deformans, infectious

CC arthritis or osteochondrosis. The present sequence is a cDNA from a

CC known gene differentially expressed in developing mesenchymal cells

XX

SQ	Sequence 2680 BP; 854 A; 446 C; 575 G; 805 T; 0 U; 0 Other;
Query Match	40.7%; Score 947; DB 6; Length 2680;
Best Local Similarity	99.5%; Pred. No. 8.8e-220;
Matches 950; Conservative	0; Mismatches 5; Indels 0; Gaps 0;
QY	1 CGGACCGTGCATATGGCCAGCGTAAAGATGCGCAAGAGCAGCGCAACAGCAGCAGCAGCG 60
DB	1 CGGACCGTGCATATGGCCAGCGTAAAGATGCGCAAGAGCAGCGCAACAGCAGCAGCAGCG 60
QY	61 GCTCCGCGCAGCGGTAGCAGCAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 120
DB	61 GCTCCGCGCAGCGGTAGCAGCAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 120
QY	121 AGCATGCGAGCAGCAGCAAGATGGGAGGAAAGCGGACTCTCGGAACTTCAATCTTCACGT 180
DB	121 AGCATGCGAGCAGCAGCAAGATGGGAGGAAAGCGGACTCTCGGAACTTCAATCTTCACGT 180
QY	181 GGTATTGCTGATGATGCAATGCTGGCGGCTCTGGAATCTGTAGCTGCTGTTGTTGATC 240
DB	181 GGTATTGCTGATGATGCAATGCTGGCGGCTCTGGAATCTGTAGCTGCTGTTGTTGATC 240
QY	241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGATGCTGATGATGATGAG 300
DB	241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGATGCTGATGATGATGAG 300
QY	301 ATTTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
DB	301 ATTTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
QY	361 CAGCAGTCCGCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTGTGG 420
DB	361 CAGCAGTCCGCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTGTGG 420
QY	421 AGCGAAGCCCGAGCAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCTCCATG 480
DB	421 AGCGAAGCCCGAGCAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCTCCATG 480
QY	481 AAATGTTACACGACAGCAATGTTGAGGAGGAGCACTTTCGCAACAGCAAGATGAGCCACG 540
DB	481 AAATGTTACACGACAGCAATGTTGAGGAGGAGCACTTTCGCAACAGCAAGATGAGCCACG 540
QY	541 GAGAACCCACACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB	541 GAGAACCCACACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY	601 AGACCTTGAACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB	601 AGACCTTGAACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY	661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720
DB	661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720
QY	721 ATTCCAGTGAACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB	721 ATTCCAGTGAACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY	781 ACCAAGTCTATGAGGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB	781 ACCAAGTCTATGAGGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY	841 CAGAAAGTAACTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 900
DB	841 CAGAAAGTAACTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 900
QY	901 AAGTAAGCATTTTCTCTGTGGAAAGACAGCAGGAAAGTACACAGGAAACAAATAG 955
DB	901 AAGTAAGCATTTTCTCTGTGGAAAGACAGCAGGAAAGTACACAGGAAACAAATAG 955
RESULT 8	
ADP24208	

ID XX ADP24208 standard; cDNA; 2680 BP.  
AC ADP24208;  
XX 18-NOV-2004 (first entry)  
DT PRO polypeptide encoding cDNA SEQ ID NO:1386.  
XX ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
XX immunosuppressive; osteopathic; antidiabetic; dermatological;  
KW antipneumatic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
KW gene therapy; immune system.  
XX Unidentified.  
OS WO2004041170-A2.  
XX 21-MAY-2004.  
XX 30-OCT-2003; 2003WO-US034312.  
XX 01-NOV-2002; 2002US-0423394P.  
XX (GETH ) GENENTECH INC.  
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;  
PI Wu TD;  
XX WPI: 2004-419628/39.  
DR P-PSDB; ADP24209.  
XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
PT renal disease, or demyelinating diseases of the central or peripheral  
PT nervous system.  
XX Claim 1; SEQ ID NO 1386; 2940pp; English.  
XX The invention relates to a novel isolated nucleic acid and the PRO  
CC polypeptide encoded by it. A protein of the invention has  
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, an autoimmune or immune-mediated dermatitis, psoriasis, an allergic  
CC disease, erythema multiforme, contact dermatitis, atopic dermatitis, food  
CC allergy, allergic rhinitis, atopic disease of the lung,  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.  
XX SQ Sequence 2680 BP; 854 A; 446 C; 575 G; 805 T; 0 U; 0 Other;  
Query Match 40.7%; Score 947; DB 13; Length 2680;  
Best Local Similarity 99.5%; Pred. No. 8.8e-220;  
Matches 950; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
1 CGGACCGTGCATGCGCCAGCGTAAGAATGCAAGAGCAGCGCAACAGCAGCAGCGCG 60

Db 1 CGGACCGTGCATGCGCCAGCGTAAGAATGCAAGAGCAGCGCAACAGCAGCAGCGCG 60  
Qy 61 GCTCCGGCAGCGGTAGCAGAGTGGGGCAGCAGCAGCCCGGGCCCGGAGAGAGACAA 120  
Db 61 GCTCCGGCAGCGGTAGCAGAGTGGGGCAGCAGCAGCCCGGGCCCGGAGAGAGACAA 120  
Qy 121 AGCATGGAGGACACAAGAAATGGGAGGAAAGCGGACTCTCGGAACTTCATCTTCAGCT 180  
Db 121 AGCATGGAGGACACAAGAAATGGGAGGAAAGCGGACTCTCAGGAACTTCATCTTCAGCT 180  
Qy 181 GGTTCATGGTATGATGCTGGCGCTCTGCACATCTGTAGCTCTCGTTGGTTGATC 240  
Db 181 GGTTCATGGTATGATGCTGGCGCTCTGCACATCTGTAGCTCTCGTTGGTTGATC 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGAG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGAG 300  
Qy 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGATTCCTGTGG 420  
Db 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGATTCCTGTGG 420  
Qy 421 AGCGAAGCCCGCAGAGATATCGAAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 480  
Db 421 AGCGAAGCCCGCAGAGATATCGAAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 480  
Qy 481 AAATGTTACACGAGAACTGTTGAGGAGAGAGACTTGCACAAAGAAAGATGACCCACAG 540  
Db 481 AAATGTTACACGAGAACTGTTGAGGAGAGAGACTTGCACAAAGAAAGATGACCCACAG 540  
Qy 541 GAGAACCAACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Qy 601 AGACCCCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCGGAGAGAG 660  
Db 601 AGACCCCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCGGAGAGAG 660  
Qy 661 CAGTTTCACAGAGCTGTAATCAGGATATGAGAGAGATGATGCTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCACAGAGCTGTAATCAGGATATGAGAGAGATGATGCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACCAAGTATGAGAGATGAAAGATGCAACCATGATACAGATGATGATGAT 780  
Db 721 ATTCCAGTGAACCAAGTATGAGAGATGAAAGATGCAACCATGATACAGATGATGATGAT 780  
Qy 781 ACCAAGTCTATGAGGAAACAGCAGTATATCACTCTAGAAATGAAGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAAACAGCAGTATATCACTCTAGAAATGAAGGATAGAAATCA 840  
Qy 841 CAGAAGTAACTGCTCCCGCTGAGGATATCTCTGAGAGATTCACAGTAAATGTTAGAG 900  
Db 841 CAGAAGTAACTGCTCCCGCTGAGGATATCTCTGAGAGATTCACAGTAAATGTTAGAG 900  
Qy 901 AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAGGAGTACCACCAAGAAACAAATAG 955  
Db 901 AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAGGAGTACCACCAAGAAACAAATAG 955  
RESULT 9  
AAZ23609  
ID AAZ23609 standard; cDNA; 2442 BP.  
XX AAZ23609;  
XX 06-JAN-2000 (first entry)  
XX Human labyrinthin cDNA.





QY	781	ACCAAGTCTATGAGGAACACGACGATATATGAACTCTAGAAATGAAGGGATAGAAATCA	840
Db	665	ACCAAGTCTATGAGGAACACGACGATATATGAACTCTAGAAATGAAGGGATAGAAATCA	724
QY	841	CAGAAGTAACTGCTCCCTCGAGGATATCTCTGAGAAGATTCACAGGTAATTTGTGAAG	900
Db	725	CAGAAGTAACTGCTCCCTCGAGGATATCTCTGAGAAGATTCACAGGTAATTTGTGAAG	784
QY	901	AAGTAAGCATTTTTCTGTGGAAGAACAGCAGGAAGTACCACCAAAACAATAG	955
Db	785	AAGTAAGCATTTTTCTGTGGAAGAACAGCAGGAAGTACCACCAATCTTTAAAG	839

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer; ss  
KW Lung cancer; cytostatic; vaccine; gene therapy; cancer; ss  
XX Lung cancer; cytostatic; vaccine; gene therapy; cancer; ss

XX  
OS  
Homo sapiens

XX  
XX  
XX

homocarbonyl.

XXXXXXX

PN  
US2002172952-A1.  
XX

PD 21-NOV-2002.  
XX

10-JUL-2001; 2001US-00902941.

XX  
PR 30-JUN-1999: 99US-00346492

PR 15-OCT-1999; 99US-00419356  
PR 17 DEC 1999; 99US-00419356

PR 17-DEC-1999; 99US-00466867.  
PR 30-DEC-1999; 99US-00476300.

PR 06-MAR-2000; 2000US-00519642.  
PR 22-MAR-2000; 2000US-00523077

22-MAR-2000; 2000US-00533077.  
10-APR-2000; 2000US-00546259.

PR 27-APR-2000; 2000US-00560406.  
PR 05-JUN-2000; 2000US-00589184.

PR 11-JUL-2000; 2000US-00614124.  
PR 29-AUG-2000; 2000US-00614124.

PR 29-AUG-2000; 2000US-00651563.  
PR 08-SEP-2000; 2000US-00658824.

PR 26-SEP-2000; 2000US-00671325.  
PR 06-OCT-2000; 2000US-00677418

PR 06-OCT-2000; 2000US-00677419.  
PR 30-OCT-2000; 2000US-00702705.

PR 13-DEC-2000; 2000US-00736457.  
PR 03-MAY-2001; 2001US-00849626.

XX  
XX  
DA (CONT ) CONTINUED

PA (CORI-) CORIXA CORP.  
XX

PI Henderson RA, Wang T, Watanabe M, Carter D, Peng Y, Durham M

FI  
XX  
Durnam M, Carter D, Fanger

DR WPI; 2003-328427/31.  
XX

New polynucleotide, useful for inhibiting development of

XX  
inhibiting development of can

Example 5; SEQID NO 1791; 82p

The invention describes an is

sequences, complement or degenerate sequences, useful for preparing a complementary DNA library.

for treating or inhibiting de-

sequence represents a polynomial method for the therapy and d

Sequence 2442 BP: 781  $\lambda$ : 381

sequence 2442 BP; 181 A; 391 C







Db	605	ATTCCAGTGAACCAAGTGTAGTGAAGATGAAGATGGACCATGATACAGATGATGAACAT	664
Qy	781	ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA	840
Db	665	ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA	724
Qy	841	CAGAAGTAATGTCCTCCCTGAGGATTAATCTCTAGAGATTCACAGGTAATGTAGAAG	900
Db	725	CAGAAGTAATGTCCTCCCTGAGGATTAATCTCTAGAGATTCACAGGTAATGTAGAAG	784
Qy	901	AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGATACACACAGAAAACAAATAG	955
Db	785	AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGATACACACAGATACTTAAAG	839
RESULT 15			
ACA92448			
ID	ACA92448 standard; DNA; 2648 BP.		
XX	ACA92448;		
AC			
XX	15-JUL-2003 (first entry)		
DT			
XX	DNA encoding human PMM-33.		
DE			
XX	Human; protein modification and maintenance molecule; PMM; cancer;		
KW	cell proliferation disorder; atherosclerosis; neurological disorder;		
KW	epilepsy; Huntington's disease; stroke; immune disorder; allergy;		
KW	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;		
KW	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;		
KW	infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;		
KW	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnery;		
KW	antiinflammatory; thyromimetic; gene; ds.		
XX			
OS	Homo sapiens.		
XX			
FN	WO2003031939-A2.		
XX			
PD	17-APR-2003.		
XX			
PF	11-OCT-2002; 2002WO-US032850.		
XX			
PR	12-OCT-2001; 2001US-0329689P.		
PR	25-OCT-2001; 2001US-0335703P.		
PR	09-NOV-2001; 2001US-0348887P.		
PR	28-NOV-2001; 2001US-0334145P.		
PR	06-DEC-2001; 2001US-0337451P.		
PR	14-DEC-2001; 2001US-0340584P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;		
PI	Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;		
PI	Sprague WW, Hafalia AJA, Chawla NK, Lehr-Watson PM, Kable AE, Yue H;		
PI	Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;		
PI	Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;		
XX			
XX	WPI; 2003-430274/40.		
DR	P-PSDB; AB092053.		
XX			
PT	New human protein modification and maintenance molecules (PMM), useful		
PT	for diagnosing, treating and preventing diseases or conditions associated		
PT	with the aberrant PMM expression e.g. cancer, atherosclerosis, or		
PT	infections.		
XX			
PS	Claim 5; Page 306; 311pp; English.		
XX			
CC	The present invention relates to the isolation of human protein		
CC	modification and maintenance molecules (PMM), and the polynucleotide		
CC	sequences encoding them. A total of 40 PMM polypeptides (designated PMM		
CC	-1 to PMM-40) are disclosed. The sequences of the invention are useful		
CC	for diagnosing a condition or disease associated with the expression of		
CC	PMM in a subject, preparing a polyclonal or monoclonal antibody, and		

generating an expression profile of a sample containing the			
polynucleotides. The diseases or conditions associated with decreased			
expression or overexpression of PMM are cell proliferation disorders			
(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,			
Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,			
allergies), developmental disorders (e.g. hypothyroidism, Cushing's			
syndrome), gastrointestinal or epithelial disorders, and infections. The			
PMM polypeptides or their fragments are useful in screening compounds			
for effectiveness as agonists or antagonists of the polypeptides, or in			
altering the expression of the target polynucleotide and compounds that			
specifically bind to, or modulate the activity of the polypeptide.			
ACA92416-ACA92455 encode the human PMM polypeptides of the invention			
XX	Sequence 2648 BP; 844 A; 420 C; 545 G; 839 T; 0 U; 0 Other;		
Qy	111	AGAGAGACAAAGCATGGAGGACACAAGAAATGGAGGAAGCGGACTTCGGAACTTCA	170
Db	128	AAAGAGACAAAGCATGGAGGACACAAGAAATGGAGGAAGCGGACTTCGGAACTTCA	187
Qy	171	TTCTTCACGTGTTTATGGTGAATTCATTGCTGGCGCTCTGGACATCTGTAGTCTCGTT	230
Db	188	TTCTTCACGTGTTTATGGTGAATTCATTGCTGGCGCTCTGGACATCTGTAGTCTCGTT	247
Qy	231	TGTTTGTATCTTGTGACTATGAGGAAGTTCTAGAAAACCTAGGAATCTATGATCTGAT	290
Db	248	TGTTTGTATCTTGTGACTATGAGGAAGTTCTAGAAAACCTAGGAATCTATGATCTGAT	307
Qy	291	GGTATGGAGATTTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	350
Db	308	GGTATGGAGATTTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	367
Qy	351	ACTTCAGAGCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAG	410
Db	368	ACTTCAGAGCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAG	427
Qy	411	GTTCTCTGTGGAGGACAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAATTCAGTCC	470
Db	428	GTTCTCTGTGGAGGACAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAATTCAGTCC	487
Qy	471	CTTCTCCATGAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	530
Db	488	CTTCTCCATGAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	547
Qy	531	GGACCCACAGGAGAACCAACAAGAGGATGATGATGATGATGATGATGATGATGATGAT	590
Db	548	GGACCCACAGGAGAACCAACAAGAGGATGATGATGATGATGATGATGATGATGATGAT	607
Qy	591	GATAGATTGAGACCCCTGGAACTTGAAGTATCTCATGAGAAACCCGAGCATAGTACCAC	650
Db	608	GATAGATTGAGACCCCTGGAACTTGAAGTATCTCATGAGAAACCCGAGCATAGTACCAC	667
Qy	651	GTGGAAGAGAGATTTTTCACAAGACTGTAAATCAGGATATGGAAGAGATGATGATGAT	710
Db	668	GTGGAAGAGATTTTTCACAAGACTGTAAATCAGGATATGGAAGAGATGATGATGAT	777
Qy	711	GAATAATCCAGATTCAGTGAACCAAGTATGAGAGATGAAGATGCAACCATGATACAGAT	770
Db	678	-----CAGATTCAGTGAACCAAGTATGAGAGATGAAGATGCAACCATGATACAGAT	730
Qy	771	GATGTAAATACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGG	830
Db	731	GATGTAAATACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGG	790
Qy	831	ATAGAAATACCAAGTAACTGCTCCCTGAGGATTAATCTCTGTAAGATTCACAGGTA	890
Db	791	ATAGAAATACCAAGTAACTGCTCCCTGAGGATTAATCTCTGTAAGATTCACAGGTA	850
Qy	891	ATTGTAGAAAGTAAAGCATTTTCTGTGGAAGAACAGCAGGAGGATGATGATGATGAT	950

Db 851 ATTGTAGAAAGTAAGCAATTTTCCTGTGGAAAGAACAGCAGGAAGTACCACCAGATACT 910  
Qy 951 AATAG 955  
Db 911 TAAAG 915

Search completed: March 25, 2005, 09:48:25  
Job time : 1164 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 12:32:18 ; Search time 1232 Seconds  
(without alignments)  
11241.467 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 CGGACCTGCAATGCGCCAG.....CTGGGAACTCTGGAGAGA 2324

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5552208 seqs, 297965951 residues  
Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	9	US-09-903-248-3
2	2324	100.0	2324	9	US-09-859-604-3
3	2324	100.0	2324	9	US-09-903-063-3
4	2324	100.0	2324	9	US-09-903-216-3
5	2324	100.0	2324	9	US-09-903-199-3
6	2324	100.0	2324	9	US-09-903-023-3
7	2324	100.0	2324	10	US-09-436-184-3
8	2319.2	99.8	5358	15	US-10-084-817-63
9	947	40.7	2680	16	US-10-096-534-37
10	947	40.7	2680	16	US-10-289-909-3
11	819	35.2	2442	9	US-09-736-457-1791

12	819	35.2	2442	9	US-09-902-941-1791	Sequence 1791, Ap
13	819	35.2	2442	9	US-09-849-626-1791	Sequence 1791, Ap
14	819	35.2	2442	14	US-10-017-754-1791	Sequence 1791, Ap
15	819	35.2	2442	16	US-10-113-872-1791	Sequence 1791, Ap
16	819	35.2	2442	17	US-10-283-017-1791	Sequence 1791, Ap
17	617.6	26.6	3110	14	US-10-198-846-9938	Sequence 9938, Ap
18	464.6	20.0	502	10	US-09-918-995-34279	Sequence 34279, A
19	416.2	17.9	660	11	US-09-969-034-3202	Sequence 3202, A
20	364.4	15.7	366	17	US-10-242-535A-25661	Sequence 25661, A
21	364.4	15.7	366	17	US-10-085-783A-25661	Sequence 25661, A
22	317	13.6	495	14	US-10-066-543-277	Sequence 277, App
23	170.6	7.3	1382	17	US-10-443-622-24	Sequence 24, Appl
24	153.2	6.6	161	16	US-10-029-386-26964	Sequence 26964, A
25	153.2	6.6	592	16	US-10-039-386-13264	Sequence 13264, A
26	135	5.8	151	17	US-10-242-535A-2707	Sequence 2707, Ap
27	135	5.8	151	17	US-10-085-783A-2707	Sequence 2707, Ap
28	127.4	5.5	583	11	US-09-969-034-3266	Sequence 3266, Ap
29	96.8	4.2	671	10	US-09-991-936-1528	Sequence 1528, Ap
30	84.2	3.6	554	13	US-10-101-487-69	Sequence 69, Appl
31	83.4	3.6	554	13	US-10-101-487-106	Sequence 106, App
32	82.8	3.6	522	13	US-10-101-487-71	Sequence 71, Appl
33	82.8	3.6	530	13	US-10-101-487-73	Sequence 73, Appl
34	78.4	3.4	720	13	US-10-101-487-76	Sequence 76, Appl
35	78.2	3.4	720	13	US-10-101-487-74	Sequence 74, Appl
36	78	3.4	450	13	US-10-027-632-2825	Sequence 2825, Ap
37	78	3.4	450	17	US-10-027-632-2825	Sequence 2825, Ap
38	74.6	3.2	17137	15	US-10-311-455-163	Sequence 163, App
39	72.2	3.1	31124	13	US-10-087-192-463	Sequence 463, App
40	71.2	3.1	3912	17	US-10-282-122A-9928	Sequence 9928, Ap
41	66	2.8	453	13	US-10-027-632-56205	Sequence 56205, A
42	66	2.8	453	13	US-10-027-632-56205	Sequence 56205, A
43	66	2.8	453	17	US-10-027-632-56205	Sequence 56205, A
44	66	2.8	453	17	US-10-027-632-56205	Sequence 56205, A
45	66	2.8	68233	17	US-10-034-650-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1  
US-09-903-248-3  
; Sequence 3, Application US/09903248  
; Patent No. US20020102263A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV5  
; CURRENT APPLICATION NUMBER: US/09/903,248  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-248-3

Query Match	100.0%	Score 2324;	DB 9;	Length 2324;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2324;	Conservative	0;	Mismatches	0;
			Indels	0;
QY	1	CGGACCTGCAATGCGCCAGCTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG	60	
Db	1	CGGACCTGCAATGCGCCAGCTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG	60	
QY	61	GCTCCGACCGGTAGACAGTGGCGGACAGCAGCCCGGGCCCGGAGAGACAA	120	
Db	61	GCTCCGACCGGTAGACAGTGGCGGACAGCAGCCCGGGCCCGGAGAGACAA	120	

121 AGCATGGAGGACCAAGAATGGGAGGAAAGCGGACTCTCGGGAACCTTCTTCAAGT 180  
121 AGCATGGAGGACCAAGAATGGGAGGAAAGCGGACTCTCGGGAACCTTCTTCAAGT 180  
181 GGTATATGGTATGTCATTTGCTGGGCGTCTGGGACATCTGTAGCTGTGTTGGTTGATC 240  
181 GGTATATGGTATGTCATTTGCTGGGCGTCTGGGACATCTGTAGCTGTGTTGGTTGATC 240  
241 TTGTTGACTATGAGGAAGTTCTAGGAACCTAGGAATCTATGATCTGATGCTGATGGAG 300  
241 TTGTTGACTATGAGGAAGTTCTAGGAACCTAGGAATCTATGATGCTGATGCTGATGGAG 300  
301 ATTTTGTATGGTATGTCATTTGCTGGGCGTCTGGGACATCTGTAGCTGTGTTGGTTGATC 360  
301 ATTTTGTATGGTATGTCATTTGCTGGGCGTCTGGGACATCTGTAGCTGTGTTGGTTGATC 360  
361 CAGCAGTCCGCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTGTGG 420  
361 CAGCAGTCCGCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTGTGG 420  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
481 AAATGGTACACGCAAGACATGTTGAGGGAGAGACTTGCACCAAGAAAGATGGACCCACAG 540  
481 AAATGGTACACGCAAGACATGTTGAGGGAGAGACTTGCACCAAGAAAGATGGACCCACAG 540  
541 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGGAGCTGATGATGATGATGATTTG 600  
541 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGGAGCTGATGATGATGATGATTTG 600  
601 AGACCTCGGAACCTGAAGTATCTCATGAAGAACCGAGCATAGTTACACGCTGGAAGAGA 660  
601 AGACCTCGGAACCTGAAGTATCTCATGAAGAACCGAGCATAGTTACACGCTGGAAGAGA 660  
661 CAGTTTCAACAAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
661 CAGTTTCAACAAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
721 ATTCCAGTGAACCAAGTATGAGATGAAGATTTGCAACCATGATACAGATGATGATACAT 780  
721 ATTCCAGTGAACCAAGTATGAGATGAAGATTTGCAACCATGATACAGATGATGATACAT 780  
781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTAGAAATGAAGGATAGAAATCA 840  
781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTAGAAATGAAGGATAGAAATCA 840  
841 CAGAAGTAACTGCTCCCTCGAGGATAATCCTGTAGAAGATTCA CAGGTAATTGTAGAAG 900  
841 CAGAAGTAACTGCTCCCTCGAGGATAATCCTGTAGAAGATTCA CAGGTAATTGTAGAAG 900  
901 AAGTAAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCA CCAAGAAACAAATAGAAAAA 960  
901 AAGTAAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCA CCAAGAAACAAATAGAAAAA 960  
961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCTTAATTAATTAAT 1020  
961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCTTAATTAATTAAT 1020  
1021 TTGATAAGACTATTAAGCTGAACTTATGATGCTGAGAAAAA CTTCCGTAAGAGGAAAAA 1080  
1021 TTGATAAGACTATTAAGCTGAACTTATGATGCTGAGAAAAA CTTCCGTAAGAGGAAAAA 1080  
1081 TTGAGGAGCAGTGAATGATTAAGACTAGTACGCAATACCT CAGAGTCCAGAG 1140  
1081 TTGAGGAGCAGTGAATGATTAAGACTAGTACGCAATACCT CAGAGTCCAGAG 1140  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAGTAAATCAGG 1200  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAGTAAATCAGG 1200

1201 TGCTACGTGGAGCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCTGCAG 1260  
1201 TGCTACGTGGAGCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCTGCAG 1260  
1261 ACCTGCTGAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
1261 ACCTGCTGAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
1321 GAGGTTCCCTGCTTACCTTCGACAGATTAGTTCAAATATTTCCCAATGATGATCTTCTTAA 1380  
1321 GAGGTTCCCTGCTTACCTTCGACAGATTAGTTCAAATATTTCCCAATGATGATCTTCTTAA 1380  
1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATTAATGACAATGCAAGAAAGTTT 1440  
1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATTAATGACAATGCAAGAAAGTTT 1440  
1441 ATCAAGAGGTGCTGAGTGTGACACCACTTAATGATGGCTTGTCTAAAGTCCATTTGGCTTCA 1500  
1441 ATCAAGAGGTGCTGAGTGTGACACCACTTAATGATGGCTTGTCTAAAGTCCATTTGGCTTCA 1500  
1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATGAAT 1560  
1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATGAAT 1560  
1561 CCGGAGATCTGCGCATGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
1561 CCGGAGATCTGCGCATGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
1621 GGGTGGGAAACAAGAGGCATATAAGTGTGTATGAGCTTGGGCAACAAGAGAGGACATTTG 1680  
1621 GGGTGGGAAACAAGAGGCATATAAGTGTGTATGAGCTTGGGCAACAAGAGAGGACATTTG 1680  
1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGTAATGAGCTGGAAGCA CAGCCTTGGTGA 1740  
1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGTAATGAGCTGGAAGCA CAGCCTTGGTGA 1740  
1741 CCCCAGAAAGAAACGGGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTAA 1800  
1741 CCCCAGAAAGAAACGGGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTAA 1800  
1801 TCCGAGATGAAGCCTTGCAGTGTATGATAAGCCAAAGGTCTTCTCCTCCCTGAGGATG 1860  
1801 TCCGAGATGAAGCCTTGCAGTGTATGATAAGCCAAAGGTCTTCTCCTCCCTGAGGATG 1860  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA 1920  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA 1920  
1921 ATGAAAATGCTGCAAGAGGACTCTTAAACCTGTACTTACTAGAAAAGTTCCCGGAGA 1980  
1921 ATGAAAATGCTGCAAGAGGACTCTTAAACCTGTACTTACTAGAAAAGTTCCCGGAGA 1980  
1981 CAACAGATGCAAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGACTCACGTGT 2040  
1981 CAACAGATGCAAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGACTCACGTGT 2040  
2041 GGGCGCACACAGGCCCCACAACTGCAAGGCTCCGAATGCACTGGGCTTGGTATTTCCCA 2100  
2041 GGGCGCACACAGGCCCCACAACTGCAAGGCTCCGAATGCACTGGGCTTGGTATTTCCCA 2100  
2101 AGAAGGCTCAAGATTCGATGTGCCAACGAGACCAGGACCTGGGAGGAAGGCAAGGTGC 2160  
2101 AGAAGGCTCAAGATTCGATGTGCCAACGAGACCAGGACCTGGGAGGAAGGCAAGGTGC 2160  
2161 TCATCTTTGATGATCTCCTTTGAGCACGAGGTATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
2161 TCATCTTTGATGATCTCCTTTGAGCACGAGGTATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
2221 TATTTCATCGTGGATGTGGCATCCGGAACTGCAACCAAGCAGAGACGCAAGCCTTCCAG 2280  
2221 TATTTCATCGTGGATGTGGCATCCGGAACTGCAACCAAGCAGAGACGCAAGCCTTCCAG 2280  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324



1741 CCCAAAGAAACGGGCTACACAGATTAGTAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db  
1741 CCCAAAGAAACGGGCTACACAGATTAGTAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Qy  
1801 TCCGAGATGAAGGCTTGGCAGTGTGATGATAAGCAAGGTCTCTTCCCTGCCTGAGGATG 1860  
Db  
1801 TCCGAGATGAAGGCTTGGCAGTGTGATGATAAGCAAGGTCTCTTCCCTGCCTGAGGATG 1860  
Qy  
1861 AAAACCTGAGGAAAAAGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGGAAGAA 1920  
Db  
1861 AAAACCTGAGGAAAAAGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGGAAGAA 1920  
Qy  
1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAGTTCCCGCAGA 1980  
Db  
1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAGTTCCCGCAGA 1980  
Qy  
1981 CAACAGGATGAGAAAGGAGCAGATCAATATTCATCATGCAACCCCGGAGCTCACTGT 2040  
Db  
1981 CAACAGGATGAGAAAGGAGCAGATCAATATTCATCATGCAACCCCGGAGCTCACTGT 2040  
Qy  
2041 GCGCGCACAGGGCCCAAACTGCGAGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
Db  
2041 GCGCGCACAGGGCCCAAACTGCGAGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
Qy  
2101 AGGAAGCTGCAAGATTGCTGTCACAGAGACAGGACCTGGGAGGAAGGCAAGGTGC 2160  
Db  
2101 AGGAAGCTGCAAGATTGCTGTCACAGAGACAGGACCTGGGAGGAAGGCAAGGTGC 2160  
Qy  
2161 TCATCTTTTGATGACTCTTTGAGCAGCAGAGGTATGGCAGGATGCCTCATCTTTCCGGCTGA 2220  
Db  
2161 TCATCTTTTGATGACTCTTTGAGCAGCAGAGGTATGGCAGGATGCCTCATCTTTCCGGCTGA 2220  
Qy  
2221 TATTATCTGATGTGGTATCGGAACTGACCACTGACACACAGCAGCAGGCTTCCAG 2280  
Db  
2221 TATTATCTGATGTGGTATCGGAACTGACCACTGACACACAGCAGCAGGCTTCCAG 2280  
Qy  
2281 CAATTTAGCATGAATTCATCAAGCTTGGGAACTCTGGAGAGA 2324  
Db  
2281 CAATTTAGCATGAATTCATCAAGCTTGGGAACTCTGGAGAGA 2324

## RESULT 3

US-09-903-063-3

; Sequence 3, Application US/09903063

; Patent No. US20020114810A1

; GENERAL INFORMATION:

; APPLICANT: Wands, Jack R.

; APPLICANT: de la Monte, Suzanne M.

; APPLICANT: Ince, Nedim

; APPLICANT: Carlson, Rolf I.

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS

; FILE REFERENCE: 21486-032 DIV3

; CURRENT APPLICATION NUMBER: US/09/903,063

; PRIOR FILING DATE: 2001-10-11

; PRIOR FILING DATE: 09/436,184

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 2324

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-903-063-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTCAATGGCCAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60

Db 1 CGGACCGTCAATGGCCAGCGTGAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60

Qy 61 GCTCCGGCAGCGGTAGCACAGTGGCGGCGAGCAGAGCCCGGGCCCGGAGAGACAA 120  
Db  
Qy 121 AGCATGGAGCACAAAGTGGAGGAAAGCGGACTCTCGGAACTTCTTCACTTCACT 180  
Db  
Qy 181 GGTATATGATGATGCAATGCTGGGCGTCTGGACATCTGTAGCTGTGCTTGGTTGATC 240  
Db  
Qy 241 TTGTTCACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGCTGATGATGGAG 300  
Db  
Qy 301 ATTTTCACTGATGATGCAAGTTTATATAGGACTTAAAGAGAGATCTACTTCAAGC 360  
Db  
Qy 361 CAGCAGTCCCGCCAGAAAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 420  
Db  
Qy 421 AGCGAAGCCCGAGAAATATCGAAGATGAAGCAAGAAACAAATTCAGTCCCTTCTCATG 480  
Db  
Qy 481 AAATGGTACACGAGAACATGTTGAGGAGAGACATTTGAAAGAGATGGAACCCACAG 540  
Db  
Qy 541 GAGAACCCACAAAGAGGATGATGATTTCTTATGCGCACTGATGATGATGATGATG 600  
Db  
Qy 601 AGACCTTGGAACTTGAAGTATCTCATGAGAAACCGGACATCTTACCAGCTGGAAGAGA 660  
Db  
Qy 661 CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAAAAATCCAG 720  
Db  
Qy 721 ATTCCAGTGAAACAGTAGTAGAAGATGCAAGATGCAACATGATACAGATGATTAACAT 780  
Db  
Qy 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACTCTAGAAAAATGAAGGATAGAAATCA 840  
Db  
Qy 841 CAGAAGTAACTGCTCCCGCTGAGGATAATCTCTAGAAATTCACAGTAAATGTAGAAG 900  
Db  
Qy 901 AAGTAAGCAATTTTCTGTTGGAAGACAGCAGGAGTACCACCAAGAAACAAATAGAAAA 960  
Db  
Qy 961 AAGTAAGCAATTTTCTGTTGGAAGACAGCAGGAGTACCACCAAGAAACAAATAGAAAA 960  
Db  
Qy 961 CAGATGATCCAGAAACAAAGTAAAGAAAAAGAGCCCTTAAATTAATTAAT 1020  
Db  
Qy 1021 TTGTAAGACTATTAAGCTGAACTTCTGATGCTGAGAAAAATCTCCGTAAGAGGAAAAA 1080  
Db  
Qy 1081 TTGAGGAGCAGTCAATGCAATTTAAAGACTAGTACCAAAATACCTCAGAGTCCAGAG 1140  
Db  
Qy 1081 TTGAGGAGCAGTCAATGCAATTTAAAGACTAGTACCAAAATACCTCAGAGTCCAGAG 1140  
Db  
Qy 1141 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGTAATGAGG 1200



Db 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGGAGGAAGTAATGAGG 1200  
Qy 1201 TGCTACGTGGAGCCATCAGACCTTACCAGAGGTGGCAGCCTACCTGATGTCCCTGCGAG 1260  
Db 1201 TGCTACGTGGAGCCATCAGACCTTACCAGAGGTGGCAGCCTACCTGATGTCCCTGCGAG 1260  
Qy 1261 ACTGCTGAAGCTGAGTTGTAAGCGTGGCTCAGACAGGCAAAATTTCTAGGTGCATATGA 1320  
Db 1261 ACTGCTGAAGCTGAGTTGTAAGCGTGGCTCAGACAGGCAAAATTTCTAGGTGCATATGA 1320  
Qy 1321 GAGGTTCCCTGTTACCTTGACAGATTTAGTTCAACTATTTCCCAATGATCTTCCTTAA 1380  
Db 1321 GAGGTTCCCTGTTACCTTGACAGATTTAGTTCAACTATTTCCCAATGATCTTCCTTAA 1380  
Qy 1381 AAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAATGCCAAAGAAAGTTT 1440  
Db 1381 AAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAATGCCAAAGAAAGTTT 1440  
Qy 1441 ATGAAGAGTGTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATATCGCTTCA 1500  
Db 1441 ATGAAGAGTGTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATATCGCTTCA 1500  
Qy 1501 TCCTGAAGGCACAGAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAT 1560  
Db 1501 TCCTGAAGGCACAGAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAT 1560  
Qy 1561 CCGGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
Db 1561 CCGGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
Qy 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACACAGAGAGCACTTTG 1680  
Db 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACACAGAGAGCACTTTG 1680  
Qy 1681 CATCTGTCTGCAAGCTCTCACTACAAATGTAATGAGCTGAAAGCAAGCCTTGGTGA 1740  
Db 1681 CATCTGTCTGCAAGCTCTCACTACAAATGTAATGAGCTGAAAGCAAGCCTTGGTGA 1740  
Qy 1741 CCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db 1741 CCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTGCAAGTGTGATGGAATAAGGCAAGAGTCTCTTCCCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTGCAAGTGTGATGGAATAAGGCAAGAGTCTCTTCCCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAAGGGGACTGAGCCAGTTTCAAGCTGTCAGCAGCAGGAGGAAGAA 1920  
Db 1861 AAAACCTGAGGAAAAGGGGACTGAGCCAGTTTCAAGCTGTCAGCAGCAGGAGGAAGAA 1920  
Qy 1921 ATGAATGCTGCAAGAGGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGGAGA 1980  
Db 1921 ATGAATGCTGCAAGAGGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGGAGA 1980  
Qy 1981 CAAAGAGATGCAAGAGGACAGATCAAAATTTCCATCATGCAACCCCGGAGACTCACGTGT 2040  
Db 1981 CAAAGAGATGCAAGAGGACAGATCAAAATTTCCATCATGCAACCCCGGAGACTCACGTGT 2040  
Qy 2041 GCGCGCACACAGGCGCCCAACAACTGCAGGCTCCGAATGCACCTGGGCTTGGTGTATCCCA 2100  
Db 2041 GCGCGCACACAGGCGCCCAACAACTGCAGGCTCCGAATGCACCTGGGCTTGGTGTATCCCA 2100  
Qy 2101 AGGAAGGCTGCAAGATTCAGTGTGCAAGAGACAGGAGCTTGGGAGGAGCAAGGTGC 2160  
Db 2101 AGGAAGGCTGCAAGATTCAGTGTGCAAGAGACAGGAGCTTGGGAGGAGCAAGGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTGAGCAGAGGTATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTTGGACAGAGGTATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
Qy 2221 TATTTCATCGTGGATGTGTGGCATCCGGAATCTGCACACACAGCAGAGACGACCTTCCAG 2280

Db 2221 TATTTCATCGTGGATGTGTGGCATCCGGAATCTGACACCAACAGCAGAGACGCGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATCAAGCTTTGGAAACTCTCGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATCAAGCTTTGGAAACTCTCGAGAGA 2324  
RESULT 4  
US-09-903-216-3  
; Sequence 3, Application US/09903216  
; Patent No. US2002011481A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV2  
; CURRENT APPLICATION NUMBER: US/09/903,216  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-216-3  
Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCCGTGCAATGGCCAGCGTAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCCGTGCAATGGCCAGCGTAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGCGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
Db 61 GCTCCGCGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
Qy 121 AGCATGAGGACACAGAGATGGAGGAAGCGGAGTCTCGGAACTTCATCTTCACGT 180  
Db 121 AGCATGAGGACACAGAGATGGAGGAAGCGGAGTCTCGGAACTTCATCTTCACGT 180  
Qy 181 GGTTCATGTTGATGTCATTTGCTGGCGTCTGGACATCTGTCGTGCTGTTTGTATC 240  
Db 181 GGTTCATGTTGATGTCATTTGCTGGCGTCTGGACATCTGTCGTGCTGTTTGTATC 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTTAGGAAACTTAGGAACTTAGGAACTTAGGAACTTAGG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTTAGGAAACTTAGGAACTTAGGAACTTAGGAACTTAGG 300  
Qy 301 ATTTTCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 301 ATTTTCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGGAGTCTCTGTGG 420  
Db 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGGAGTCTCTGTGG 420  
Qy 421 AGCGAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGCGAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGTTACACCGAGAAATCTGTCAGGAGAACTTTGCAACAGAAAGATGACCCACAG 540  
Db 481 AAATGTTACACCGAGAAATCTGTCAGGAGAACTTTGCAACAGAAAGATGACCCACAG 540  
Qy 541 GAGAACCAACACAGAGAGTGTGTTCTTATGCGGACTGATGTAGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACACAGAGAGTGTGTTCTTATGCGGACTGATGTAGATGATGATGATGATGATG 600

601 AGACCCCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAAAGGA 660  
Db  
601 AGACCCCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAAAGGA 660  
Qy  
661 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Db  
661 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Qy  
721 ATTCAGTGAACCACTAGTGAAGATGAAGATTTGCCACCATGATACAGATGATGTAACAT 780  
Db  
721 ATTCAGTGAACCACTAGTGAAGATGAAGATTTGCCACCATGATACAGATGATGTAACAT 780  
Qy  
781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAATGAAGGATGAATCA 840  
Db  
781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAATGAAGGATGAATCA 840  
Qy  
841 CAGAAGTAACTGCTCCCTCAGGATATCTCTGAGAAAGATTTACAGGTAATTTGTAAG 900  
Db  
841 CAGAAGTAACTGCTCCCTCAGGATATCTCTGAGAAAGATTTACAGGTAATTTGTAAG 900  
Qy  
901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGTACCAACAGAAACAAATAGAAAA 960  
Db  
901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGTACCAACAGAAACAAATAGAAAA 960  
Qy  
961 CAGATGATCCAGAACCAAGCTTAAGTAAGAAAGAGCCTTAACTTTTAATTAAT 1020  
Db  
961 CAGATGATCCAGAACCAAGCTTAAGTAAGAAAGAGCCTTAACTTTTAATTAAT 1020  
Qy  
1021 TTGATAAGCACTATTAAAGCTGAATCTGAGTGAAGAAATCTCGTAAAGGGGAAAA 1080  
Db  
1021 TTGATAAGCACTATTAAAGCTGAATCTGAGTGAAGAAATCTCGTAAAGGGGAAAA 1080  
Qy  
1081 TTGAGGAAGCACTGAATGAATTAAGAACTAGTACGCAATACCTCAGAGTCCAGAG 1140  
Db  
1081 TTGAGGAAGCACTGAATGAATTAAGAACTAGTACGCAATACCTCAGAGTCCAGAG 1140  
Qy  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
Db  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
Qy  
1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCTCCTGATGTCCTGAG 1260  
Db  
1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCTCCTGATGTCCTGAG 1260  
Qy  
1261 ACTGCTGAAGCTGATTTGAAGCTCGCTCAGACAGGCAACAATTTCTAGTCAATGA 1320  
Db  
1261 ACTGCTGAAGCTGATTTGAAGCTCGCTCAGACAGGCAACAATTTCTAGTCAATGA 1320  
Qy  
1321 GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTTAA 1380  
Db  
1321 GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTTAA 1380  
Qy  
1381 AAAATGACCTTGGCTGGGATACCTCTGATAGGAGATTAAGCAATGCAAGAAAGTTT 1440  
Db  
1381 AAAATGACCTTGGCTGGGATACCTCTGATAGGAGATTAAGCAATGCAAGAAAGTTT 1440  
Qy  
1441 ATGAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATATGCTTCA 1500  
Db  
1441 ATGAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATATGCTTCA 1500  
Qy  
1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGAAATAGAT 1560  
Db  
1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGAAATAGAT 1560  
Qy  
1561 CCGGAGATCCTGCACTGATGAGGAGATTTTATTTCCACCTGGGGATGCCATGCAGA 1620  
Db  
1561 CCGGAGATCCTGCACTGATGAGGAGATTTTATTTCCACCTGGGGATGCCATGCAGA 1620  
Qy  
1621 GGGTTGGGAACAAAGAGGATATAAGTGTGATGAGCTTGGGCACAGAGAGGACATTTG 1680  
Db  
1621 GGGTTGGGAACAAAGAGGATATAAGTGTGATGAGCTTGGGCACAGAGAGGACATTTG 1680

Qy 1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGGAATGAAAGCAGAGCTTGGTGA 1740  
Db 1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGGAATGAAAGCAGAGCTTGGTGA 1740  
Qy 1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
Db 1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTCAGTGTGATGGAATAAGCCAAAGGTCCTTCTCTGCTCAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTCAGTGTGATGGAATAAGCCAAAGGTCCTTCTCTGCTCAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAGAAAGGGGACTGGAGCCAGTTTCAAGTGTGGCAGCAAGAAAGAA 1920  
Db 1861 AAAACCTGAGGAAAGAAAGGGGACTGGAGCCAGTTTCAAGTGTGGCAGCAAGAAAGAA 1920  
Qy 1921 ATGAAAATCCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTTCCCGAGA 1980  
Db 1921 ATGAAAATCCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTTCCCGAGA 1980  
Qy 1981 CAAACAGATGCAAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGACTCAGTGT 2040  
Db 1981 CAAACAGATGCAAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGACTCAGTGT 2040  
Qy 2041 GGCCGCACACAGGCGCCCAAAAAGTGGAGCTCGAATGCACTGGCTTGGTGAATTTCCCA 2100  
Db 2041 GGCCGCACACAGGCGCCCAAAAAGTGGAGCTCGAATGCACTGGCTTGGTGAATTTCCCA 2100  
Qy 2101 AGGAAGCTGCAAGATTTCCGATGTGCCAACGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
Db 2101 AGGAAGCTGCAAGATTTCCGATGTGCCAACGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
Qy 2161 TCATCTTTGATGATCTCTTTGAGCAGAGATGAGCAGATGCTCATCTTTCCGCTGA 2220  
Db 2161 TCATCTTTGATGATCTCTTTGAGCAGAGATGAGCAGATGCTCATCTTTCCGCTGA 2220  
Qy 2221 TATTCATCTGGAATGTGGCATCCGGAATGACACACAGCAGAGACGAGCTTCCAG 2280  
Db 2221 TATTCATCTGGAATGTGGCATCCGGAATGACACACAGCAGAGACGAGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAATCTCGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAATCTCGGAGAGA 2324

RESULT 5  
US-09-903-199-3  
; Sequence 3, Application US/09903199  
; Patent No. US20020122802A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV4  
; CURRENT APPLICATION NUMBER: US/09/903,199  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-199-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGAATGGCCAGCGTGAAGATGCCAAGAGACGCGGCAACAGCAGCAGCGG 60



Qy	2221	TATTATCGTGGATGTGTGCGATCCGAACTGTACACACAGCAGACGAGCCTTCCAG	2280
Db	2221	TATTATCGTGGATGTGTGCGATCCGAACTGTACACACAGCAGACGAGCCTTCCAG	2280
Qy	2281	CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA	2324
Db	2281	CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA	2324
RESULT 6			
US-09-903-023-3			
; Sequence 3, Application US/09903023			
; Patent No. US20020146421A1			
; GENERAL INFORMATION:			
; APPLICANT: Wands, Jack R.			
; APPLICANT: de la Monte, Suzanne M.			
; APPLICANT: Ince, Nedim			
; APPLICANT: Carlson, Rolf I.			
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS			
; FILE REFERENCE: 21486-032 DIV1			
; CURRENT APPLICATION NUMBER: US/09/903,023			
; CURRENT FILING DATE: 2001-10-11			
; PRIOR APPLICATION NUMBER: 09/436,184			
; PRIOR FILING DATE: 1999-11-08			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 3			
; LENGTH: 2324			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-903-023-3			
Query Match 100.0%; Score 2324; DB 9; Length 2324;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CGGACCGTGCAATGCCCCAGCGTAAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG	60
Db	1	CGGACCGTGCAATGCCCCAGCGTAAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG	60
Qy	61	GCTCCGGCAGCGGTAGCAGCGTGGCGGAGCAGCAGCCCCGGGGCCCGGAGAGAGACAA	120
Db	61	GCTCCGGCAGCGGTAGCAGCGTGGCGGAGCAGCAGCCCCGGGGCCCGGAGAGAGACAA	120
Qy	121	AGCATGGAGGACACAAGAATGGGAGGAAAGCGGACTCTCGGAAACTTCATCTTCAAGT	180
Db	121	AGCATGGAGGACACAAGAATGGGAGGAAAGCGGACTCTCGGAAACTTCATCTTCAAGT	180
Qy	181	GGTTTATGGTGAATGCTGGCGGTCTGCACATCTGTAGCTGTGCTTTCGTTTGTTGATC	240
Db	181	GGTTTATGGTGAATGCTATGCTGGCGGTCTGCACATCTGTAGCTGTGCTTTCGTTTGTTGATC	240
Qy	241	TTGTTTGACTATCAGGAAGTTCTTAGGAAAACTAGGAATCTATGATGCTGATGGTGGAG	300
Db	241	TTGTTTGACTATCAGGAAGTTCTTAGGAAAACTAGGAATCTATGATGCTGATGGTGGAG	300
Qy	301	ATTTTGATGTGATGATGCAAAAGTTTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360
Db	301	ATTTTGATGTGATGATGCAAAAGTTTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360
Qy	361	CAGCAGTCCCGCAGAGAGAGGCTGAGCCACACACTCAGCCGAGGAGCAGGTTCTCTGTGG	420
Db	361	CAGCAGTCCCGCAGAGAGAGGCTGAGCCACACACTCAGCCGAGGAGCAGGTTCTCTGTGG	420
Qy	421	AGGCAAGACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG	480
Db	421	AGGCAAGACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG	480
Qy	481	AAATGTTACACGACAGACATGTTGAGGGAAGACTTGGCAACAAGAGATGGACCCACAG	540
Db	481	AAATGTTACACGACAGACATGTTGAGGGAAGACTTGGCAACAAGAGATGGACCCACAG	540

Db 1621 GGGTTGGGAACAAGAGGCATATAGTGGTATGAGCTTGGGCACAGAGAGGACACTTTG 1680  
 Qy 1681 CATCTGTCTGGCAACGCTCACTTACAATGTGAATGGACTGAAAGCAGAGCCTTGGTGGGA 1740  
 Db 1681 CATCTGTCTGGCAACGCTCACTTACAATGTGAATGGACTGAAAGCAGAGCCTTGGTGGGA 1740  
 Qy 1741 CCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAA 1800  
 Db 1741 CCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAA 1800  
 Qy 1801 TCCGAGATGAAGCCCTTCCAGTGATGATAAAGCCAAAGCTCTTCCCTGCTGAGGATG 1860  
 Db 1801 TCCGAGATGAAGCCCTTCCAGTGATGATAAAGCCAAAGCTCTTCCCTGCTGAGGATG 1860  
 Qy 1861 AAAAATCTGAGGGAAGAGGAGCTGGAGCCAGTTACGCTGTGGCAGCAAGGAAGAA 1920  
 Db 1861 AAAAATCTGAGGGAAGAGGAGCTGGAGCCAGTTACGCTGTGGCAGCAAGGAAGAA 1920  
 Qy 1921 ATGAAATGCTTGCAGAGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCCGAGA 1980  
 Db 1921 ATGAAATGCTTGCAGAGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCCGAGA 1980  
 Qy 1981 CAACAGGATGCAAGAGAGGAGCAGATCAAAATATTCCATCATGCAACCCGGGACTCAGCTG 2040  
 Db 1981 CAACAGGATGCAAGAGAGGAGCAGATCAAAATATTCCATCATGCAACCCGGGACTCAGCTG 2040  
 Qy 2041 GGCCGCACACAGGCCCCACAACTGCAGGCTCGGAATCGCACTGGGCTTGGTGAATCCCA 2100  
 Db 2041 GGCCGCACACAGGCCCCACAACTGCAGGCTCGGAATCGCACTGGGCTTGGTGAATCCCA 2100  
 Qy 2101 AGGAAGGCTGCAAGATTGCGATGTGCCAAGCAGACCAAGCACTGGGAGGAAGGCAAGGTGC 2160  
 Db 2101 AGGAAGGCTGCAAGATTGCGATGTGCCAAGCAGACCAAGCACTGGGAGGAAGGCAAGGTGC 2160  
 Qy 2161 TCATCTTTGATGACTCTTTTGACACAGGATATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
 Db 2161 TCATCTTTGATGACTCTTTTGACACAGGATATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
 Qy 2221 TATTTCATCGTGGATGTGGCATCCGGAACTGACACCAAGCAGGAGCAGGACCTTCAG 2280  
 Db 2221 TATTTCATCGTGGATGTGGCATCCGGAACTGACACCAAGCAGGAGCAGGACCTTCAG 2280  
 Qy 2281 CAATTTAGCATGAATTCAGCAGCTTGGGAACTCTGGAGAGA 2324  
 Db 2281 CAATTTAGCATGAATTCAGCAGCTTGGGAACTCTGGAGAGA 2324

RESULT 7

US-09-436-184-3  
 ; Sequence 3, Application US/09436184  
 ; Publication No. US20030031670A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wands, Jack R.  
 ; APPLICANT: de la Monte, Suzanne M.  
 ; APPLICANT: Ince, Nedim  
 ; APPLICANT: Carlson, Rolf I.  
 ; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
 ; FILE REFERENCE: R.I. Hosp. Malignant Neoplasms  
 ; CURRENT APPLICATION NUMBER: US/09/436,184  
 ; CURRENT FILING DATE: 1999-11-08  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 2324  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-436-184-3

Query Match 100.0%; Score 2324; DB 10; Length 2324;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCATATGCCCCAGCGTAAGATGCTCAAGAGCAGCGGCAACAGACGACGACGCG 60  
 Db 1 CGGACCGTGCATATGCCCCAGCGTAAGATGCTCAAGAGCAGCGGCAACAGACGACGACGCG 60  
 Qy 61 GCTCCGGCAGCGGTAGCACAGTTCGGGCGAGCAGAGCCCGGGGCGCGGAGAGACAA 120  
 Db 61 GCTCCGGCAGCGGTAGCACAGTTCGGGCGAGCAGAGCCCGGGGCGCGGAGAGACAA 120  
 Qy 121 AGCATGGAGGACACAAGATGGAGAAAGGCGGACTCTCGGGAACCTTCACTTCTTCACT 180  
 Db 121 AGCATGGAGGACACAAGATGGAGAAAGGCGGACTCTCGGGAACCTTCACTTCTTCACT 180  
 Qy 181 GGTATTATGGTGAATTCGATTCGGGCGTCTGGACATCTGTAGCTGTGCTGCTGCTGCTG 240  
 Db 181 GGTATTATGGTGAATTCGATTCGGGCGTCTGGACATCTGTAGCTGTGCTGCTGCTGCTG 240  
 Qy 241 TTGTTGACTATGAGGAAGTCTTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300  
 Db 241 TTGTTGACTATGAGGAAGTCTTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300  
 Qy 301 ATTTTGTGATGGATGATGCAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
 Db 301 ATTTTGTGATGGATGATGCAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
 Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCCGAGGAGCAGGTTCCCTGTGG 420  
 Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCCGAGGAGCAGGTTCCCTGTGG 420  
 Qy 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAGAAAGCAAAATTCAGTCCCTTCCATG 480  
 Db 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAGAAAGCAAAATTCAGTCCCTTCCATG 480  
 Qy 481 AAATGGTACACCGCAGACATGTTGAGGGAAGACTTTCGCAAGAAAGATGACCCACAG 540  
 Db 481 AAATGGTACACCGCAGACATGTTGAGGGAAGACTTTCGCAAGAAAGATGACCCACAG 540  
 Qy 541 GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATGATG 600  
 Db 541 GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATGATG 600  
 Qy 601 AGACCTCTGGAACCTCGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAAGA 660  
 Db 601 AGACCTCTGGAACCTCGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAAGA 660  
 Qy 661 CAGTTTACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
 Db 661 CAGTTTACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
 Qy 721 ATTCCAGTGAAACCACTAGTAGAAGATGAAAGATGCAACCATGATACAGATGATGTAACAT 780  
 Db 721 ATTCCAGTGAAACCACTAGTAGAAGATGAAAGATGCAACCATGATACAGATGATGTAACAT 780  
 Qy 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
 Db 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
 Qy 841 CAGAAGTAACTCTCCCTCCTGAGGATATCTCTGAGAGATTCACAGGTAATCTGTAAG 900  
 Db 841 CAGAAGTAACTCTCCCTCCTGAGGATATCTCTGAGAGATTCACAGGTAATCTGTAAG 900  
 Qy 901 AAGTAAGCATTTTCTGCTGGAGAGACAGCAGGAGTACCACAGAAACAAATAGAAAAA 960  
 Db 901 AAGTAAGCATTTTCTGCTGGAGAGACAGCAGGAGTACCACAGAAACAAATAGAAAAA 960  
 Qy 961 CAGATGATCCAGAACAAAAAGCAAAAGTTTAAAGAAAAAGAGCCTTAAATTAAT 1020  
 Db 961 CAGATGATCCAGAACAAAAAGCAAAAGTTTAAAGAAAAAGAGCCTTAAATTAAT 1020  
 Qy 1021 TTGATAAGACTATTAAAGCTGAACTTGTGCTGCAGAAAAAACTCCGTAAGAGGGGAAAA 1080  
 Db 1021 TTGATAAGACTATTAAAGCTGAACTTGTGCTGCAGAAAAAACTCCGTAAGAGGGGAAAA 1080  
 Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCCAAATACCTTCAGAGTCCAGAG 1140

1081 |TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCTAGAGTCCAGAG 1140  
1141 |CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
1141 |CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
1201 |TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCGAG 1260  
1201 |TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCGAG 1260  
1261 |ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGCAAAATTTCTAGGTCATATGA 1320  
1261 |ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGCAAAATTTCTAGGTCATATGA 1320  
1321 |GAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTCACATTTTCCCAATGATCTCTCTAA 1380  
1321 |GAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTCACATTTTCCCAATGATCTCTCTAA 1380  
1381 |AAATGACCTTGGCGTGGATACCTCTTGATAGAGATATGACAATGCAAGAAAGTTT 1440  
1381 |AAATGACCTTGGCGTGGATACCTCTTGATAGAGATATGACAATGCAAGAAAGTTT 1440  
1441 |ATGAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTTAAAGTCCATTTATGGCTCA 1500  
1441 |ATGAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTTAAAGTCCATTTATGGCTCA 1500  
1501 |TCCTGAAGGCACAGAAACAAATTTGCTGAGAGATCCCATATTTAAAGAGGAATAGAT 1560  
1501 |TCCTGAAGGCACAGAAACAAATTTGCTGAGAGATCCCATATTTAAAGAGGAATAGAT 1560  
1561 |CCGGAGATCTGGCACTGATGATGGAGATTTATTTCCACTGGGGATGCCATGAGA 1620  
1561 |CCGGAGATCTGGCACTGATGATGGAGATTTATTTCCACTGGGGATGCCATGAGA 1620  
1621 |GGGTTGGGAAACAAAGGCGATATAAGTGTATGATGCTTTGGGCACAAAGAGAGACATTTG 1680  
1621 |GGGTTGGGAAACAAAGGCGATATAAGTGTATGATGCTTTGGGCACAAAGAGAGACATTTG 1680  
1681 |CATCTGCTGGCAACGCTCACTCTCAATGTGAATGGACTGAAGACACAGCCTTGGTGA 1740  
1681 |CATCTGCTGGCAACGCTCACTCTCAATGTGAATGGACTGAAGACACAGCCTTGGTGA 1740  
1741 |CCCCAAAGAAACGGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAATTA 1800  
1741 |CCCCAAAGAAACGGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAATTA 1800  
1801 |TCCGAGATGAAGGCTTGCAGTGAATGATTAAGCCAAAGGTCTCTTCTGCTGAGGATG 1860  
1801 |TCCGAGATGAAGGCTTGCAGTGAATGATTAAGCCAAAGGTCTCTTCTGCTGAGGATG 1860  
1861 |AAAACTGAGGAAAAAGGGGACTGAGCCAGTTCACTGCTGGCAGCAAGGAAGAA 1920  
1861 |AAAACTGAGGAAAAAGGGGACTGAGCCAGTTCACTGCTGGCAGCAAGGAAGAA 1920  
1921 |ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
1921 |ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
1981 |CAACAGGATGCAAGAGGACAGATCAATATTTCCATCATGCAACCCGGGACTCAGGTG 2040  
1981 |CAACAGGATGCAAGAGGACAGATCAATATTTCCATCATGCAACCCGGGACTCAGGTG 2040  
2041 |GGCGGCACAGAGGCCCCAACAATGCAAGGCTCCGAATGCACTTGGGCTTGGTATTTCCA 2100  
2041 |GGCGGCACAGAGGCCCCAACAATGCAAGGCTCCGAATGCACTTGGGCTTGGTATTTCCA 2100  
2101 |AGGAAGGCTGCAAGATTCGATGTGCCAAACAGACACAGGACTGGGAGGAGCAAGGTGC 2160  
2101 |AGGAAGGCTGCAAGATTCGATGTGCCAAACAGACACAGGACTGGGAGGAGCAAGGTGC 2160  
2161 |TCATCTTTGATGATCTCTTTTGGACACGAGGTATGGCAGGATGCTCATCTTTTCCGGCTGA 2220

2161 |TCATCTTTGATGATCTCTTTTGGACACGAGGTATGGCAGGATGCTCATCTTTTCCGGCTGA 2220  
2221 |TATTTCATCTGATGCTGTGGCATCCGGAACCTGACACCAACAGCAGACGCGACCTTCCAG 2280  
2221 |TATTTCATCTGATGCTGTGGCATCCGGAACCTGACACCAACAGCAGACGCGACCTTCCAG 2280  
2281 |CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324  
2281 |CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324

RESULT 8  
US-10-084-817-63  
; Sequence 63, Application US/10084817  
; Publication No. US20030119009A1  
; GENERAL INFORMATION:  
; APPLICANT: Susan Stuart  
; APPLICANT: Jed G. Nuchtern  
; APPLICANT: Sharon E. Pion  
; APPLICANT: Jason M. Shohet  
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
; FILE REFERENCE: PA-0046 US  
; CURRENT APPLICATION NUMBER: US/10/084,817  
; CURRENT FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/270,784  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 365  
; SOFTWARE: PERL Program  
; SEQ ID NO 63  
; LENGTH: 5358  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030119009A1 089172.13  
US-10-084-817-63

Query Match 99.8%; Score 2319.2; DB 15; Length 5358;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGGACCTGCAATGGCCCGCCGCTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 258 CGGACCTGCAATGGCCCGCCGCTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 317  
QY 61 GCTCCGCGCAGCGGTAGCAGCGAGTGGGGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
DB 318 GCTCCGCGCAGCGGTAGCAGCGAGTGGGGCAGCAGCGCCCGGGCCCGGAGAGACAA 377  
QY 121 AGCATGAGGACACAAAGATGGGAGGAAAGCGGACTCTCGGAACTTCAATTTTCACGT 180  
DB 378 AGCATGAGGACACAAAGATGGGAGGAAAGCGGACTCTCGGAACTTCAATTTTCACGT 437  
QY 181 GGTTTATGTTGATTTGCAATTTGCTGGCGCTCTGGCAATCTGTAGCTGTGTTGTTGATC 240  
DB 438 GGTTTATGTTGATTTGCAATTTGCTGGCGCTCTGGCAATCTGTAGCTGTGTTGTTGATC 497  
QY 241 TTGTTGATGAGCAAGTCTTAGGAAACTTAGGAATCTATGATGCTGATGTTGATGAG 300  
DB 498 TTGTTGATGAGCAAGTCTTAGGAAACTTAGGAATCTATGATGCTGATGTTGATGAG 557  
QY 301 ATTTTGTGATGATGTCGCAAAAGTCTTTATTTAGGACTTAAAGAGAGATCTATTTTCAAGC 360  
DB 558 ATTTTGTGATGATGTCGCAAAAGTCTTTATTTAGGACTTAAAGAGAGATCTATTTTCAAGC 617  
QY 361 CAGCAGTCCCGCAGAGAGCTGAGCACAACCTGAGCCCGGAGCAGGCTTCTCTGG 420  
DB 618 CAGCAGTCCCGCAGAGAGCTGAGCACAACCTGAGCCCGGAGCAGGCTTCTCTGG 677  
QY 421 AGGCAGAACCCAGCAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
DB 678 AGGCAGAACCCAGCAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 737

QY 481 AATGTGACACGACGACATGTTGAGGGAGAGACTTGCAACAAGAGATGACCCACAG 540  
DB 738 AATGTGTACACGACGACATGTTGAGGGAGAGACTTGCAACAAGAGATGACCCACAG 797  
QY 541 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGAGTCTGATGATGATGATGATG 600  
DB 798 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGAGTCTGATGATGATGATGATG 857  
QY 601 AGACCTGTCGACCTGAGTATCTCATGAAGAACCGAGCATAGTTTACACGTGGAAGAGA 660  
DB 858 AGACCTGTCGACCTGAGTATCTCATGAAGAACCGAGCATAGTTTACACGTGGAAGAGA 917  
QY 661 CAGTTTCACAGACTGTAAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
DB 918 CAGTTTCACAGACTGTAAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 977  
QY 721 ATTCCAGTGAACCAAGTATGAGAGATGGAAGATGTCACCATGATACAGATGATTAACAT 780  
DB 978 ATTCCAGTGAACCAAGTATGAGAGATGGAAGATGTCACCATGATACAGATGATTAACAT 1037  
QY 781 ACCAAGTCTATGAGAACCAAGCAGTATATGAACCTCTAGAAATGAAAGGATGAAATCA 840  
DB 1038 ACCAAGTCTATGAGAACCAAGCAGTATATGAACCTCTAGAAATGAAAGGATGAAATCA 1097  
QY 841 CAGAGTAACTGCTCCCTCAGGATATCTCTGTAAGAGATTCACAGGTAATTTAGAGAG 900  
DB 1098 CAGAGTAACTGCTCCCTCAGGATATCTCTGTAAGAGATTCACAGGTAATTTAGAGAG 1157  
QY 901 AAGTAACTGCTCCCTCAGGATATCTCTGTAAGAGATTCACAGGTAATTTAGAGAG 960  
DB 1158 AAGTAACTGCTCCCTCAGGATATCTCTGTAAGAGATTCACAGGTAATTTAGAGAG 1217  
QY 961 CAGATGATCCAGAACCAAGGCAAAAGTTAAGAAAGAGAGCTTAACTTTAAATTAAT 1277  
DB 1218 CAGATGATCCAGAACCAAGGCAAAAGTTAAGAAAGAGAGCTTAACTTTAAATTAAT 1337  
QY 1021 TTGATAGACTATTAAGCTGACCTGATGCTGCGAAGAACTCCGTAAGAGGGAAGAA 1080  
DB 1278 TTGATAGACTATTAAGCTGACCTGATGCTGCGAAGAACTCCGTAAGAGGGAAGAA 1337  
QY 1081 TTGAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
DB 1338 TTGAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1397  
QY 1141 CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGATGATGAGG 1200  
DB 1398 CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGATGATGAGG 1457  
QY 1201 TGCTACGTGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCGTCAG 1260  
DB 1458 TGCTACGTGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCGTCAG 1517  
QY 1261 ACCTGCTGAGCTGAGTTTGAAGCTGCTCAGACAGGCAACAATTTCTAGTGCATATGA 1320  
DB 1518 ACCTGCTGAGCTGAGTTTGAAGCTGCTCAGACAGGCAACAATTTCTAGTGCATATGA 1577  
QY 1321 GAGTTCCCTGCTTACCTGCGAGAGATGATTTCAACTATTTCCCAATGATATCTTCTTAA 1380  
DB 1578 GAGTTCCCTGCTTACCTGCGAGAGATGATTTCAACTATTTCCCAATGATATCTTCTTAA 1637  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATATGACATGCAAGAGAGTTT 1440  
DB 1638 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATATGACATGCAAGAGAGTTT 1597  
QY 1441 ATGAAGAGGTGCTGAGTGTGACCTAATGATGCTTGTGTAAGTCCATTTATGCTTCA 1500  
DB 1698 ATGAAGAGGTGCTGAGTGTGACCTAATGATGCTTGTGTAAGTCCATTTATGCTTCA 1757  
QY 1501 TCCTGAAGGCAAGAACCAAAATTCCTGAGAGCATGCCATATTTAAAGGAAGAGATGAAT 1560  
DB 1758 TCCTGAAGGCAAGAACCAAAATTCCTGAGAGCATGCCATATTTAAAGGAAGAGATGAAT 1817  
QY 1561 CCGGAGATCCTGGCAGTATGAGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620

DB 1818 CCGGAGATCCTGGCAGTATGAGAGATTTTATTTCCACCTGGGGATGCCATGCAGA 1877  
QY 1621 GGGTTGGGAAACAAGAGGCAATATAGTGTATGAGCTTGGGCAACAAGAGAGGACACTTTG 1680  
DB 1878 GGGTTGGGAAACAAGAGGCAATATAGTGTATGAGCTTGGGCAACAAGAGAGGACACTTTG 1937  
QY 1681 CATCTGTCTGCGCAACGCTCACTCTACAATGTGAATGAGCTGAAAGCACACGCTTGTGGA 1740  
DB 1938 CATCTGTCTGCGCAACGCTCACTCTACAATGTGAATGAGCTGAAAGCACACGCTTGTGGA 1997  
QY 1741 CCCCACAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
DB 1998 CCCCACAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 2057  
QY 1801 TCCGAGATGAAGGCTTGGCAGTGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
DB 2058 TCCGAGATGAAGGCTTGGCAGTGATGATGATGATGATGATGATGATGATGATGATGATG 2117  
QY 1861 AAAACCTGAGGGGAAAGGGGACTGGAGCCAGTTTACGCTGTGGCAGCAAGGAAGAA 1920  
DB 2118 AAAACCTGAGGGGAAAGGGGACTGGAGCCAGTTTACGCTGTGGCAGCAAGGAAGAA 2177  
QY 1921 ATGAAAATGCTCTCAAAAGGAGCTCTTAAACCTCTGTAACCTTACTAGAAAAGTTCCCGGAGA 1980  
DB 2178 ATGAAAATGCTCTCAAAAGGAGCTCTTAAACCTCTGTAACCTTACTAGAAAAGTTCCCGGAGA 2237  
QY 1981 CAACAGGATGACAGAGAGGACAGATCAATATTTCCATCATGCAACCCCGGACTCAGCTGT 2040  
DB 2238 CAACAGGATGACAGAGAGGACAGATCAATATTTCCATCATGCAACCCCGGACTCAGCTGT 2297  
QY 2041 GGGCGCACACAGGGGCCCAAACTGCAAGCTTCGAGCTTCGAAATGCACTGGGCTTGGTATCCCA 2100  
DB 2298 GGGCGCACACAGGGGCCCAAACTGCAAGCTTCGAGCTTCGAAATGCACTGGGCTTGGTATCCCA 2357  
QY 2101 AGGAAGCTGCAAGATTCGATGTGCAACAGAGACAGAGACCTGGGAGGAAGCAAGGTGC 2160  
DB 2358 AGGAAGCTGCAAGATTCGATGTGCAACAGAGACAGAGACCTGGGAGGAAGCAAGGTGC 2417  
QY 2161 TCATCTTGTATGACTCTCTTGGACACAGGATGATGAGGAGATGCTCATCTTTCGGCTGA 2220  
DB 2418 TCATCTTGTATGACTCTCTTGGACACAGGATGATGAGGAGATGCTCATCTTTCGGCTGA 2477  
QY 2221 TATTCATCTGAGTGTGCTGCTTGGATCGGATCGGACAGGATGATGAGGAGATGCTCATCTTTCGGCTGA 2280  
DB 2478 TATTCATCTGAGTGTGCTGCTTGGATCGGATCGGACAGGATGATGAGGAGATGCTCATCTTTCGGCTGA 2537  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324  
DB 2538 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2581

## RESULT 9

US-10-096-534-37  
; Sequence 37, Application US/10096534  
; Publication No. US20030166887A1  
; GENERAL INFORMATION:  
; APPLICANT: The Brigham and Women's Hospital, Inc.  
; APPLICANT: Yates, Karen  
; APPLICANT: Mizuno, Shuichi  
; APPLICANT: Glowacki, Julie  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF SKELETAL DEGENERATION CONDITIONS  
; FILE REFERENCE: B0801/7244/KA/ERP  
; CURRENT APPLICATION NUMBER: US/10/096,534  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US 60/274,980  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 37  
; LENGTH: 2680  
; TYPE: DNA  
; ORGANISM: Homo sapiens



US-10-096-534-37

Query Match 40.7%; Score 947; DB 16; Length 2680;  
Best Local Similarity 99.5%; Pred. No. 1.3e-214;  
Matches 950; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATGCGCGGAGTAAAGTCCAGAGAGCGGCGCAACAGCAGCAGCAGCG 60  
DB 1 CGGACCGTGCATGCGCGGAGTAAAGTCCAGAGAGCGGCGCAACAGCAGCAGCAGCG 60

QY 61 GCTCCGCGAGCGGTAGCACAGTGCAGGCGAGCAGCGCGCGCGCGCGCGAGAGACAA 120  
DB 61 GCTCCGCGAGCGGTAGCACAGTGCAGGCGAGCAGCGCGCGCGCGCGAGAGACAA 120

QY 121 AGCATGAGGACACAAAGATGGGAGGAAAGCGGACTCTCAGGAATCTTCACTTCACT 180  
DB 121 AGCATGAGGACACAAAGATGGGAGGAAAGCGGACTCTCAGGAATCTTCACTTCACT 180

QY 181 GGTTCATGCTGATGCTGCGGCGTCTGGACATCTGTAGCTGTCTTGTGTTGATC 240  
DB 181 GGTTCATGCTGATGCTGCGGCGTCTGGACATCTGTAGCTGTCTTGTGTTGATC 240

QY 241 TTGTTGACTATGAGGAGTCTTAGGAAAATCTAGGAATCTATGATGCTGATGTTGATG 300  
DB 241 TTGTTGACTATGAGGAGTCTTAGGAAAATCTAGGAATCTATGATGCTGATGTTGATG 300

QY 301 ATTTTGTGATGATGCTGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACT 360  
DB 301 ATTTTGTGATGATGCTGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACT 360

QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCGCGCGAGGAGGAGTCTTCTG 420  
DB 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCGCGCGAGGAGGAGTCTTCTG 420

QY 421 AGGAGAGACCCAGAGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACT 480  
DB 421 AGGAGAGACCCAGAGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACT 480

QY 481 AATGTTGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACTTCACT 540  
DB 481 AATGTTGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACTTCACT 540

QY 541 GAGAACCCAGAGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACTTCACT 600  
DB 541 GAGAACCCAGAGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACTTCACT 600

QY 601 AGACCCCTGGAACCTGGAAGTATCTCAAGAAACCCGAGCATAGTTACCACTGGAAGAG 660  
DB 601 AGACCCCTGGAACCTGGAAGTATCTCAAGAAACCCGAGCATAGTTACCACTGGAAGAG 660

QY 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGATGATGCTGAGCAGGAAATCCAG 720  
DB 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGATGATGCTGAGCAGGAAATCCAG 720

QY 721 ATTCCAGTGAACCAAGTATGAGGAGTAAAGTCTTAAAGAGAGATGATGATGATGAT 780  
DB 721 ATTCCAGTGAACCAAGTATGAGGAGTAAAGTCTTAAAGAGAGATGATGATGATGAT 780

QY 781 ACCAGTCTATGAGGAGCAGGATATGGAAGTCTTAAAGAGAGATGATGATGATGAT 840  
DB 781 ACCAGTCTATGAGGAGCAGGATATGGAAGTCTTAAAGAGAGATGATGATGATGAT 840

QY 841 CAGAACTGATGCTCCCTGAGGATTAATCTTGTAGAGATTAATCAGGATTAATGTAAG 900  
DB 841 CAGAACTGATGCTCCCTGAGGATTAATCTTGTAGAGATTAATCAGGATTAATGTAAG 900

QY 901 AAGTAAAGCATTTTCTGTTGAGAGACAGCAGGAGTACCACCAAGAAATAG 955  
DB 901 AAGTAAAGCATTTTCTGTTGAGAGACAGCAGGAGTACCACCAAGAAATAG 955

Sequence 3, Application US/10269909  
Publication No. US20030180747A1

GENERAL INFORMATION:  
APPLICANT: HRUBAN, RALPH H.  
APPLICANT: ARGANI, PEDRAM  
APPLICANT: IACOBUIO-DONAHUE, CHRISTINE  
APPLICANT: MAITRA, ANIRBAN  
TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES  
FILE REFERENCE: 58303(71699)  
CURRENT APPLICATION NUMBER: US/10/269,909  
CURRENT FILING DATE: 2003-10-11  
PRIOR APPLICATION NUMBER: 60/328,609  
PRIOR FILING DATE: 2001-10-11  
PRIOR APPLICATION NUMBER: 60/332,754  
PRIOR FILING DATE: 2001-11-19  
NUMBER OF SEQ ID NOS: 87  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 2680  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-269-909-3

Query Match 40.7%; Score 947; DB 16; Length 2680;  
Best Local Similarity 99.5%; Pred. No. 1.3e-214;  
Matches 950; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATGCGCGGAGTAAAGTCCAGAGAGCGGCGCAACAGCAGCAGCAGCG 60  
DB 1 CGGACCGTGCATGCGCGGAGTAAAGTCCAGAGAGCGGCGCAACAGCAGCAGCAGCG 60

QY 61 GCTCCGCGAGCGGTAGCACAGTGCAGGCGAGCAGCGCGCGCGCGCGAGAGACAA 120  
DB 61 GCTCCGCGAGCGGTAGCACAGTGCAGGCGAGCAGCGCGCGCGCGCGAGAGACAA 120

QY 121 AGCATGAGGACACAAAGATGGGAGGAAAGCGGACTCTCAGGAATCTTCACTTCACT 180  
DB 121 AGCATGAGGACACAAAGATGGGAGGAAAGCGGACTCTCAGGAATCTTCACTTCACT 180

QY 181 GGTTCATGCTGATGCTGCGGCGTCTGGACATCTGTAGCTGTCTTGTGTTGATC 240  
DB 181 GGTTCATGCTGATGCTGCGGCGTCTGGACATCTGTAGCTGTCTTGTGTTGATC 240

QY 241 TTGTTGACTATGAGGAGTCTTAGGAAAATCTAGGAATCTATGATGCTGATGTTGATG 300  
DB 241 TTGTTGACTATGAGGAGTCTTAGGAAAATCTAGGAATCTATGATGCTGATGTTGATG 300

QY 301 ATTTTGTGATGATGCTGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACT 360  
DB 301 ATTTTGTGATGATGCTGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACT 360

QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCGCGCGAGGAGTCTTCTG 420  
DB 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCGCGCGAGGAGTCTTCTG 420

QY 421 AGGAGAGACCCAGAGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACT 480  
DB 421 AGGAGAGACCCAGAGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACT 480

QY 481 AATGTTGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACTTCACT 540  
DB 481 AATGTTGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACTTCACT 540

QY 541 GAGAACCCAGAGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACTTCACT 600  
DB 541 GAGAACCCAGAGATGAGGAGTAAAGTCTTAAAGAGAGATCTTCACTTCACTTCACT 600

QY 601 AGACCCCTGGAACCTGGAAGTATCTCAAGAAACCCGAGCATAGTTACCACTGGAAGAG 660  
DB 601 AGACCCCTGGAACCTGGAAGTATCTCAAGAAACCCGAGCATAGTTACCACTGGAAGAG 660

QY 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGATGATGCTGAGCAGGAAATCCAG 720  
DB 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGATGATGCTGAGCAGGAAATCCAG 720

Db 661 CAGTTTCAAGACAGTGTAAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCAACATGATACAGATGATGTAACAT 780  
Db 721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCAACATGATACAGATGATGTAACAT 780  
Qy 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTGAAATGAAAGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTGAAATGAAAGGATAGAAATCA 840  
Qy 841 CAGAAGTAACTCTCCCTCCCTGAGGATATCTGTGTGAAGATTCACAGGTAATTTGTAAG 900  
Db 841 CAGAAGTAACTCTCCCTCCCTGAGGATATCTGTGTGAAGATTCACAGGTAATTTGTAAG 900  
Qy 901 AAGTAAGCAATTTTCTGTGTGAAGAAACAGCAGGAAAGTACCACAGAAACAAATAG 955  
Db 901 AAGTAAGCAATTTTCTGTGTGAAGAAACAGCAGGAAAGTACCACAGATACTTTAAAG 955

## RESULT 11

US-09-736-457-1791  
; Sequence 1791, Application US/09736457  
; Patent No. US20020168637A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1791

Query Match 35.2%; Score 819; DB 9; Length 2442;  
Best Local Similarity 98.8%; Pred. No. 3.7e-184;  
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 121 AGCATGAGACACAGATGGGAAAGCGGACCTCGGAACTTCATTTCCAGT 180  
Db 5 AGCTTGAAGACACAGATGGGAAAGCGGACCTCGGAACTTCATTTCCAGT 64  
Qy 181 GGTATTGTTGATGCTGGCGCTCGACATCTGAGCTGCTGTTGGTTGATC 240  
Db 65 GGTATTGTTGATGCTGGCGCTCGACATCTGAGCTGCTGTTGGTTGATC 124  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGAACTATGATGCTGATGATGGAG 300  
Db 125 TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGAACTATGATGCTGATGATGGAG 184  
Qy 301 ATTTTGTATGATGATGCTAGGATTTTATTAGGACTTAAAGAGATCTACTTCAGGC 360  
Db 185 ATTTTGTATGATGATGCTAGGATTTTATTAGGACTTAAAGAGATCTACTTCAGGC 244  
Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACATGAGCCCGAGGAGCAGGTTCTGTGG 420  
Db 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACATGAGCCCGAGGAGCAGGTTCTGTGG 304  
Qy 421 AGGCAGAACCCAGATATCGAAGATGAAGCAAAAGAAATTCAGTCCCTTCTCCATG 480

Db 305 AGCGAGAACCCAGCAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 364  
Qy 481 AAATGGTACACGCGAAGACATGTTGAGGGAGAAAGACTTGCACAAAGAAAGATGGACCCACAG 540  
Db 365 AAATGGTACACGCGAAGACATGTTGAGGGAGAAAGACTTGCACAAAGAAAGATGGACCCACAG 424  
Qy 541 GAGAACCAACCAACAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTTG 600  
Db 425 GAGAACCAACCAACAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTTG 484  
Qy 601 AGACCTCGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACGTCGGAAGCA 660  
Db 485 AGACCTCGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACGTCGGAAGCA 544  
Qy 661 CAGTTTCCAGAGACTCTGAATCAGGATATGGAAGATGATGTCTGAGCAGGAAATTCAG 720  
Db 545 CAGTTTCCAGAGACTCTGAATCAGGATATGGAAGATGATGTCTGAGCAGGAAATTCAG 604  
Qy 721 ATTCAGTGAACCCAGTAGTAGAAGATGAAGATTCACCATGATACAGATGATGTAACAT 780  
Db 605 ATTCAGTGAACCCAGTAGTAGAAGATGAAGATTCACCATGATACAGATGATGTAACAT 664  
Qy 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840  
Db 665 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 724  
Qy 841 CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAGAGATTCACAGGTAATTTAGAG 900  
Db 725 CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAGAGATTCACAGGTAATTTAGAG 784  
Qy 901 AAGTAAGCAATTTTCTGTGTGAAGAAACAGCAGGAAAGTACCACAGAAACAAATAG 955  
Db 785 AAGTAAGCAATTTTCTGTGTGAAGAAACAGCAGGAAAGTACCACAGATACTTTAAAG 839

## RESULT 12

US-09-902-941-1791  
; Sequence 1791, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Mainerakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-902-941-1791

Query Match 35.2%; Score 819; DB 9; Length 2442;  
Best Local Similarity 98.8%; Pred. No. 3.7e-184;  
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 121 AGCATGAGACACAGATGGGAAAGCGGACCTCGGAACTTCATTTCCAGT 180  
Db 5 AGCTTGAAGACACAGATGGGAAAGCGGACCTCGGAACTTCATTTCCAGT 64  
Qy 181 GGTATTGTTGATGCTGGCGCTCGACATCTGAGCTGCTGTTGGTTGATC 240





Mon Mar 28 06:06:59 2005

Qy	661	CAGTTTCACAAGACTGTAAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAAATCCAG	720
Db	545		604
Qy	721	ATTCCAGTGAAACCAGTAGTAGAAGATCAAGATTGCACCATGATACAGATGATGTAACAT	780
Db	605		664
Qy	781	ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAAATCA	840
Db	665		724
Qy	841	CAGAAGTAACTGTCTCCCTGAGGATTAATCCTGTAGAGATTCCACAGGTAATTGTAGAAG	900
Db	725		784
Qy	901	AAGTAAGCATTTCCTGTGGAAGAACAGCAGGGAAGTACCACAGAAACAAATAG	955
Db	785		839

Search completed: March 25, 2005, 17:33:28  
Job time : 1237 secs



QY	361	CAGCAGTCCGCCAGAAAGAGCGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCCTGTGG	420	QY	1441	ATGAAGAGGTCTCAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA	1500
Db	361			Db	1441		1500
QY	421	AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACTTCAAGTCCCTTCTCCATG	480	QY	1501	TCCTGAAGGCACAGAACAAATTCCTGAGAGCATCCCATATTTAAAGGAAGGATAGAAAT	1560
Db	421			Db	1501		1560
QY	481	AAATGGTACACGACAGACATGTTGAGGGAGAAGACTTGCACCAAGAGATGGACCCACAG	540	QY	1561	CGGAGATCCTGGCACTGATGATGGGAGATTTATTTTCCACTGGGGATGCCATGCGA	1620
Db	481			Db	1561		1620
QY	541	GAGAACCAACAAAGAGATGATGAGTTTCTTATGGCGACTGATGATGATGATGATGATG	600	QY	1621	GGGTTGGGAAACAAAGAGGCATATAGTGGTATGAGCTTTGGGGCAAGAGAGAGACACTTTG	1680
Db	541			Db	1621		1680
QY	601	AGACCTGGAACTGAACTATCTCATGAAGAAACCGAGCATAGTTACACGCTGGGAAGAGA	660	QY	1681	CATCTGTCTGGCAACGCTCACCTACTAATGTGAATGGACTGAAAGCACAAGCCTTGGTGA	1740
Db	601			Db	1681		1740
QY	661	CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG	720	QY	1741	CCCCAAAAGAAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAGAAGAACTGGAGTTAA	1800
Db	661			Db	1741		1800
QY	721	ATTCAGTGAACCACTAGTAGAAGTGAAGATTTGCACCATGATACAGATGATGTAACAT	780	QY	1801	TCGAGATGAAGCCCTTTCAGTGTATGATGATGATGATGATGATGATGATGATGATG	1860
Db	721			Db	1801		1860
QY	781	ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAAGGAGATGAATCA	840	QY	1861	AAAACTGAGGAAAAGGGGACTGGAGCCAGTTTCAAGTGTGGCAGCAAGAAAGAA	1920
Db	781			Db	1861		1920
QY	841	CAGAGTAACTGCTCCCTGAGGATATCTGTAGAGATTTCAAGTAACTTTTGAAG	900	QY	1921	ATGAAAATGCTGCAAGAGGACTCTTAAACCTGTACTTACTAGAAAAGTTCCCGGAGA	1980
Db	841			Db	1921		1980
QY	901	AGTAAAGCAATTTTCTGTGGAAGAAACAGCAGGAAGTACCACCAACAAATAGAAAA	960	QY	1981	CAACAGATGCAAGAGGACAGATCAAAATATTTCAATCATGCAACCCCGGAGCTCAGTGT	2040
Db	901			Db	1981		2040
QY	961	CAGATGATCAGAACAAAGCAAAAGTTAAGAAAAGAGCCTTAACCTTTTAATAAT	1020	QY	2041	GGCGGCACACAGGGCCCAACAACTGAGGCTCGAATGCACTGGGCTTGGTGAITCCCA	2100
Db	961			Db	2041		2100
QY	1021	TTGATGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1080	QY	2101	AGGAAGCTCAGAGATTCGATGTCGCAACGAGACAGGACCTGGGAGGAGCAAGGTGC	2160
Db	1021			Db	2101		2160
QY	1081	TTGAGGAAGCAGTGAATGCAATTTAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAG	1140	QY	2161	TCATCTTTGATGACTCCTTTTGAAGCAGAGGATGCGAGGATGCTCATCTTTCCGGCTGA	2220
Db	1081			Db	2161		2220
QY	1141	CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTCAGAGGAGGAGTAAATGAGG	1200	QY	2221	TATTCATCGTGGATGTGGCATCCGGAACTGACACACAGCAGAGACGCGCTTCCAG	2280
Db	1141			Db	2221		2280
QY	1201	TGCTAGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCCTACCTGATGTCCCTGCAG	1260	QY	2281	CAATTTAGCTGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2324
Db	1201			Db	2281		2324
QY	1261	ACCTGCTGAAAGCTGAGTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGTGCATATCA	1320				
Db	1261						
QY	1321	GAGTTCCTGCTTACCTCAGAGATAGTTCAACTATTTTCCCAATGATCTTCCCTTAA	1380				
Db	1321						
QY	1381	AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATATGCAATGCAAGAAAGTTT	1440				
Db	1381						

RESULT 2  
US-09-903-199-3  
; Sequence 3, Application US/09903199  
; Patent No. 6797696  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV4  
; CURRENT APPLICATION NUMBER: US/09/903,199

; CURRENT FILING DATE: 2001-07-11  
; PRIOR FILING DATE: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-199-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATGGCCAGCGTAAGAATGCCAAGAGCAGCGCAACAGCAGCAGCG 60  
DB 1 CGGACCGTGCATGGCCAGCGTAAGAATGCCAAGAGCAGCGCAACAGCAGCAGCG 60

QY 61 GCTCCGCGAGCGGTAGCACAGTGCAGGAGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
DB 61 GCTCCGCGAGCGGTAGCACAGTGCAGGAGCAGCAGCGCCCGGGCCCGGAGAGACAA 120

QY 121 AGCATGGAGGACACAAAGATGGAGGAAAGCGGAGCTCTCGGAACTTCAATCTTCACGT 180  
DB 121 AGCATGGAGGACACAAAGATGGAGGAAAGCGGAGCTCTCGGAACTTCAATCTTCACGT 180

QY 181 GGTATTATGGTGAATGCTGGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC 240  
DB 181 GGTATTATGGTGAATGCTGGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC 240

QY 241 TTGTTGACTATGAGGAGTCTAGGAAATAGGAATCTATGATGCTGATGAGTGGAG 300  
DB 241 TTGTTGACTATGAGGAGTCTAGGAAATAGGAATCTATGATGCTGATGAGTGGAG 300

QY 301 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
DB 301 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360

QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGTTCTGTGG 420  
DB 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGTTCTGTGG 420

QY 421 AGGAGAACCCAGAAATATCGAAGATGAAGCAAGAAACAAATTCAGTCCCTTCCATG 480  
DB 421 AGGAGAACCCAGAAATATCGAAGATGAAGCAAGAAACAAATTCAGTCCCTTCCATG 480

QY 481 AAATGGTACACGACGACATGTTGAGGAGAGAGACTTGCACCAAGAGATGGACCCACAG 540  
DB 481 AAATGGTACACGACGACATGTTGAGGAGAGAGACTTGCACCAAGAGATGGACCCACAG 540

QY 541 GAGAACCCACCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
DB 541 GAGAACCCACCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

QY 601 AGACCTCGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAGAGA 660  
DB 601 AGACCTCGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAGAGA 660

QY 661 CAGTTTCAAGACTGTAATCAGGATATGAGAGATGATGATGATGATGATGATGATGATG 720  
DB 661 CAGTTTCAAGACTGTAATCAGGATATGAGAGATGATGATGATGATGATGATGATGATG 720

QY 721 ATTCAGTGAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
DB 721 ATTCAGTGAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780

QY 781 ACCAAGTCTATGAGGAAACAGCAGATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
DB 781 ACCAAGTCTATGAGGAAACAGCAGATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840

QY 841 CAGAAGTAACTGCTCCCTCAGGATAATCTGTAGAGATTCACAGGTAATTTGTAGAG 900  
DB 841 CAGAAGTAACTGCTCCCTCAGGATAATCTGTAGAGATTCACAGGTAATTTGTAGAG 900

DB 841 CAGAAGTAACTGCTCCCTCAGGATAATCTGTAGAGATTCACAGGTAATTTGTAGAG 900

QY 901 AAGTAAGCATTTTCTGTGGAAGAAACAGCAGAAAGTACCAACAGAAACAAATAGAAAA 960  
DB 901 AAGTAAGCATTTTCTGTGGAAGAAACAGCAGAAAGTACCAACAGAAACAAATAGAAAA 960

QY 961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTAAATCTTTTAAATAAAT 1020  
DB 961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTAAATCTTTTAAATAAAT 1020

QY 1021 TTGATAAGACTATTAAAGCTGAACTTGTGCTGCAGAAAAAATCCGTAAAGGGGAAAA 1080  
DB 1021 TTGATAAGACTATTAAAGCTGAACTTGTGCTGCAGAAAAAATCCGTAAAGGGGAAAA 1080

QY 1081 TTGAGGAGCAGTGAATGCAATTTAAGAACTAGTAGCAATACCTCAGAGTCCACGAG 1140  
DB 1081 TTGAGGAGCAGTGAATGCAATTTAAGAACTAGTAGCAATACCTCAGAGTCCACGAG 1140

QY 1141 CAGATATGGAAAGGCGCAGTGTGAGATGATTTGGCTGAGAGAGAGAAATGAGG 1200  
DB 1141 CAGATATGGAAAGGCGCAGTGTGAGATGATTTGGCTGAGAGAGAGAAATGAGG 1200

QY 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGTCCTGCAG 1260  
DB 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGTCCTGCAG 1260

QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAATTTCTAGGTCAATGA 1320  
DB 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAATTTCTAGGTCAATGA 1320

QY 1321 GAGGTTCCTGCTTACCTGCGAGAGATTAGTTCAACTATTTCCCAATGATCTCTCTTAA 1380  
DB 1321 GAGGTTCCTGCTTACCTGCGAGAGATTAGTTCAACTATTTCCCAATGATCTCTCTTAA 1380

QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
DB 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440

QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTCA 1500  
DB 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTCA 1500

QY 1501 TCCTGAAGGACACAGAAACAAATTTGCTGAGCATCCCATATTTTAAAGAGGAGATAGAT 1560  
DB 1501 TCCTGAAGGACACAGAAACAAATTTGCTGAGCATCCCATATTTTAAAGAGGAGATAGAT 1560

QY 1561 CCGAGATCTCTGGCACTGATGATGGGAGATTTTATTTCCACTGGGGGATGCGCATGCGAGA 1620  
DB 1561 CCGAGATCTCTGGCACTGATGATGGGAGATTTTATTTTCCACTGGGGGATGCGCATGCGAGA 1620

QY 1621 GGGTTGGGAAACAAAGAGGCGATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACATTTG 1680  
DB 1621 GGGTTGGGAAACAAAGAGGCGATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACATTTG 1680

QY 1681 CATCTGCTCGCAACCGCTCACTCTACATGTAATGTAATGTAATGTAATGTAATGTAATG 1740  
DB 1681 CATCTGCTCGCAACCGCTCACTCTACATGTAATGTAATGTAATGTAATGTAATGTAATG 1740

QY 1741 CCCCAAGAGAAACGGGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
DB 1741 CCCCAAGAGAAACGGGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800

QY 1801 TCCGAGATGAAGCCCTTGCAGTGGATTAAGGCAAAAGGTCTCTTCTGCTCGCTGAGGATG 1860  
DB 1801 TCCGAGATGAAGCCCTTGCAGTGGATTAAGGCAAAAGGTCTCTTCTGCTCGCTGAGGATG 1860

QY 1861 AAAACCTGAGGGAAGGAGGAGCTGGAGCCAGTTTCAAGCTGTGGCAGAGAGGAAAGAA 1920  
DB 1861 AAAACCTGAGGGAAGGAGGAGCTGGAGCCAGTTTCAAGCTGTGGCAGAGAGGAAAGAA 1920

QY 1921 ATGAAAATGCTCAGAGGAGCTCTTAAACCTGCTTACTAGTAAAGTTTCCCGAGA 1980  
DB 1921 ATGAAAATGCTCAGAGGAGCTCTTAAACCTGCTTACTAGTAAAGTTTCCCGAGA 1980



QY 1981 CAACAGGATGAGAGGAGGACAGATCAAAATATTCATCATGCAACCCCGGACTCACGTGT 2040  
Db 1981 CAACAGGATGAGAGGAGGACAGATCAAAATATTCATCATGCAACCCCGGACTCACGTGT 2040  
QY 2041 GGCGCACACAGGCGCCACAAACTGCGAGTCCGAATGCACTGGGCTTGGTATCCCA 2100  
Db 2041 GGCGCACACAGGCGCCACAAACTGCGAGTCCGAATGCACTGGGCTTGGTATCCCA 2100  
QY 2101 AGAAGGTCGAAGATTCGATGTCGCAACAGAGACAGGACCTGGGAGGAGGCAAGTGC 2160  
Db 2101 AGAAGGTCGAAGATTCGATGTCGCAACAGAGACAGGACCTGGGAGGAGGCAAGTGC 2160  
QY 2161 TCATCTTTGATGACTCCTTTGAGCAGCAGGATGATGCGAGGATGCTCATCTTTCCGGCTGA 2220  
Db 2161 TCATCTTTGATGACTCCTTTGAGCAGCAGGATGATGCGAGGATGCTCATCTTTCCGGCTGA 2220  
QY 2221 TATTCATGCTGGATGTGTGGCATCCGGAATGACACACAGCAGAGACGAGCCTCCAG 2280  
Db 2221 TATTCATGCTGGATGTGTGGCATCCGGAATGACACACAGCAGAGACGAGCCTCCAG 2280  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 3  
US-09-903-216-3  
; Sequence 3, Application US/09903216  
; Patent No. 6812206  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV2  
; CURRENT APPLICATION NUMBER: US/09/903,216  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-216-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCGTGCAATGCGCCAGCGTGAAGATGCCAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCAATGCGCCAGCGTGAAGATGCCAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
QY 61 GCTCCGCGAGCGGTAGCAGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120  
Db 61 GCTCCGCGAGCGGTAGCAGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120  
QY 121 AGCATGGAGGACACAAGAAATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
Db 121 AGCATGGAGGACACAAGAAATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
QY 181 GGTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
Db 181 GGTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
QY 241 TTGTTGACTATCAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGATGATG 300  
Db 241 TTGTTGACTATCAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGATGATG 300

QY 301 ATTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 301 ATTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCCTGTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCCTGTGG 420  
QY 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
QY 481 AANTGTGACACGAGCAACATGTTGGAGGAGAGACACTTGCACACAGAGATGAGCCACAG 540  
Db 481 AANTGTGACACGAGCAACATGTTGGAGGAGAGACTTGCACACAGAGATGAGCCACAG 540  
QY 541 GAGAACCAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
QY 601 AGACCCCTGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCACGTGGAAGA 660  
Db 601 AGACCCCTGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCACGTGGAAGA 660  
QY 661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGAT 720  
Db 661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGAT 720  
QY 721 ATTCCAGTGAACCAAGTAGTAGAAGATGCAACCATGATGATGATGATGATGATGATGATG 780  
Db 721 ATTCCAGTGAACCAAGTAGTAGAAGATGCAACCATGATGATGATGATGATGATGATGATG 780  
QY 781 ACCAAGTCTATGAGGAAACAGCAGATATGAACTCTAGAAAAATGAAGGATGAAGATCA 840  
Db 781 ACCAAGTCTATGAGGAAACAGCAGATATGAACTCTAGAAAAATGAAGGATGAAGATCA 840  
QY 841 CAGAGTAACTGCTCCCTGAGGATTAATCTCTAGAGATTTCCAGGATTAATTTAGTAGAG 900  
Db 841 CAGAGTAACTGCTCCCTGAGGATTAATCTCTAGAGATTTCCAGGATTAATTTAGTAGAG 900  
QY 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGTACCACAGAAACAAATAGAAAAA 960  
Db 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGTACCACAGAAACAAATAGAAAAA 960  
QY 961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATTAAT 1020  
Db 961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATTAAT 1020  
QY 1021 TTGATAAGACTATTAAGCTGAACTTGATGCTGCAGAAAAAATCCCGTAAAGGGGAAAAA 1080  
Db 1021 TTGATAAGACTATTAAGCTGAACTTGATGCTGCAGAAAAAATCCCGTAAAGGGGAAAAA 1080  
QY 1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTTCAGAGTCCACGAG 1140  
QY 1141 CAAGATATGGAAGCCGAGTGTGAGATGATTTGGCTGAGAGAGAGAGATTAATCAGG 1200  
Db 1141 CAAGATATGGAAGCCGAGTGTGAGATGATTTGGCTGAGAGAGAGAGATTAATCAGG 1200  
QY 1201 TGCTACCTGGAGCCATCGAGACCTTACCAGAGGTGGCCAGCTTACCTGATGTCCTCGAG 1260  
Db 1201 TGCTACCTGGAGCCATCGAGACCTTACCAGAGGTGGCCAGCTTACCTGATGTCCTCGAG 1260  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCAATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCAATGA 1320  
QY 1321 GAGTTCCCTGCTTACCTGAGAGATAGTTCAACTTTTCCCAATGATCTTCCCTTAA 1380  
Db 1321 GAGTTCCCTGCTTACCTGAGAGATAGTTCAACTTTTCCCAATGATCTTCCCTTAA 1380  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGATGAATGACAAATGCAAGAAAGTTT 1440

Db 1381 AAAATGACCTTGGCTGGATACCTCTTGATAGAGATAATGACAAATGCAAGAAATTT 1440  
Qy 1441 ATGAAGAGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCAATATGCTTCA 1500  
Db 1441 ATGAAGAGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCAATATGCTTCA 1500  
Qy 1501 TCTGAAGCCACAGAAATGCTGAGAGCATCCATATTTAAAGGAGGATAGAT 1560  
Db 1501 TCTGAAGCCACAGAAATGCTGAGAGCATCCATATTTAAAGGAGGATAGAT 1560  
Qy 1561 CCGGAGATCTGGCACTGATGAGGAGATTTTATTCACCTGGGGATGCCATCAGA 1620  
Db 1561 CCGGAGATCTGGCACTGATGAGGAGATTTTATTCACCTGGGGATGCCATCAGA 1620  
Qy 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACAAAGAGGACATTTG 1680  
Db 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACAAAGAGGACATTTG 1680  
Qy 1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGAGCTTGAAGAGGACACCTTTG 1740  
Db 1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGAGCTTGAAGAGGACACCTTTG 1740  
Qy 1741 CCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db 1741 CCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1860  
Db 1801 TCCGAGATGAAGGCTTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1860  
Qy 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTCAAGCTGTGGCAGCAAGGAAGAA 1920  
Db 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTCAAGCTGTGGCAGCAAGGAAGAA 1920  
Qy 1921 ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTGACCTTACTAGAAAGTTCCCGGAG 1980  
Db 1921 ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTGACCTTACTAGAAAGTTCCCGGAG 1980  
Qy 1981 CAACAGATGCAAGAGAGAGAGAGATCAATATTCATCATGACCCCGGAGATCAAGTGT 2040  
Db 1981 CAACAGATGCAAGAGAGAGAGAGATCAATATTCATCATGACCCCGGAGATCAAGTGT 2040  
Qy 2041 GGGCGCACAGGGCCCAAACTGCAAGGCTCCGAATGACCTGGGCTTGGTGAATCCCA 2100  
Db 2041 GGGCGCACAGGGCCCAAACTGCAAGGCTCCGAATGACCTGGGCTTGGTGAATCCCA 2100  
Qy 2101 AGGAAGGCTGCAAGATTCGATGTGCAACGAGACGAGGCTGGGAGGAGCAAGGTGC 2160  
Db 2101 AGGAAGGCTGCAAGATTCGATGTGCAACGAGACGAGGCTGGGAGGAGCAAGGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGGGAGGATGGGAGGATGGGAGGATGG 2220  
Db 2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGGGAGGATGGGAGGATGGGAGGATGG 2220  
Qy 2221 TATTCATGCTGAGTGTGGGATCCGGAATGACACACAGCAGAGAGCGACCTTCCAG 2280  
Db 2221 TATTCATGCTGAGTGTGGGATCCGGAATGACACACAGCAGAGAGCGACCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 4

US-09-903-063-3  
; Sequence 3, Application US/09903063  
; Patent No. 6815415  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV3  
; CURRENT APPLICATION NUMBER: US/09/903,063  
; CURRENT FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-063-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCTGCAATGGCCGAGCGTAAGAATGCCAAGAGAGCGGCAACAGCAGCAGCG 60  
Db 1 CGGACCTGCAATGGCCGAGCGTAAGAATGCCAAGAGAGCGGCAACAGCAGCAGCG 60  
Qy 61 GCTCCGCGAGCGGTAGCAGCGAGTGGCGCAGCAGCAGCCCCGGGCGCCGAGAGAGACA 120  
Db 61 GCTCCGCGAGCGGTAGCAGCGAGTGGCGCAGCAGCAGCCCCGGGCGCCGAGAGAGACA 120  
Qy 121 AGCATGGAGGACACAAGAAATGGAGGAAAGCGGACTCTCGGGAACCTTCCTTCACGT 180  
Db 121 AGCATGGAGGACACAAGAAATGGAGGAAAGCGGACTCTCGGGAACCTTCCTTCACGT 180  
Qy 181 GGTATTGCTGATGCTGCTGCTGGCGCTGCGACATCTGTAGCTGTGCTGCTGCTGCTG 240  
Db 181 GGTATTGCTGATGCTGCTGCTGGCGCTGCGACATCTGTAGCTGTGCTGCTGCTGCTG 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACCTAGGAATCTATGATGCTGATGCTGATG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAACCTAGGAATCTATGATGCTGATGCTGATG 300  
Qy 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Qy 361 CAGCAGTCCGCGCAGAGAGCTCAGCCACACACTGAGCCGCGAGAGAGAGTTCCTGTGG 420  
Db 361 CAGCAGTCCGCGCAGAGAGCTCAGCCACACACTGAGCCGCGAGAGAGAGTTCCTGTGG 420  
Qy 421 AGCAGAACCCAGAAATTCGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
Db 421 AGCAGAACCCAGAAATTCGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
Qy 481 AAATGTTACACGAG 540  
Db 481 AAATGTTACACGAG 540  
Qy 541 GAGAACCAACAG 600  
Db 541 GAGAACCAACAG 600  
Qy 601 AGACCTCGGACCTGAGAGATCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
Db 601 AGACCTCGGACCTGAGAGATCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
Qy 661 CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Db 661 CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Qy 721 ATTCCAGTGAACAG 780  
Db 721 ATTCCAGTGAACAG 780  
Qy 781 ACCAAGTCTATGAGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
Db 781 ACCAAGTCTATGAGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840

841 CAGAAGTAACTGCTCCCTGAGGATAATCTCTAGAGATTCAAGGTAAATTTAGAG 900  
841 CAGAAGTAACTGCTCCCTGAGGATAATCTCTAGAGATTCAAGGTAAATTTAGAG 900  
901 AAGTAAGCATTTTCTGTGGAAGAACAGCAAGTACCACAGAAAACAAATAGAAAA 960  
901 AAGTAAGCATTTTCTGTGGAAGAACAGCAAGTACCACAGAAAACAAATAGAAAA 960  
961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTAAATTTTAAATTAAT 1020  
961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTAAATTTTAAATTAAT 1020  
1021 TTGTAAGACTATTAAAGCTGAATTTGATGCTGAGAAAACCTCCGTAAAGGGGAAAA 1080  
1021 TTGTAAGACTATTAAAGCTGAATTTGATGCTGAGAAAACCTCCGTAAAGGGGAAAA 1080  
1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCTCAGAGTCCACGAG 1140  
1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCTCAGAGTCCACGAG 1140  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAAATAGAGG 1200  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAAATAGAGG 1200  
1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTTACCTGATGCTCCGTCAG 1260  
1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTTACCTGATGCTCCGTCAG 1260  
1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320  
1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320  
1321 GAGGTTCCTGCTTACCTCGCAGAGATAGTTCACCTATTTCCCAATGATCTTCCCTTAA 1380  
1321 GAGGTTCCTGCTTACCTCGCAGAGATAGTTCACCTATTTCCCAATGATCTTCCCTTAA 1380  
1381 AAAATGACCTTGGCTGGGATACCTCTTGTAGAGAGATAATGACAAATGCAAGAAAGTTT 1440  
1381 AAAATGACCTTGGCTGGGATACCTCTTGTAGAGAGATAATGACAAATGCAAGAAAGTTT 1440  
1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTTAGGCTTCA 1500  
1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTTAGGCTTCA 1500  
1501 TCCTGAAGCCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGAAAGAAATAGAT 1560  
1501 TCCTGAAGCCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGAAAGAAATAGAT 1560  
1561 CCGGAGATCCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
1561 CCGGAGATCCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
1621 GGGTTGGGAAACAAAGAGGCATTAAGTGTATGAGCTTTGGGCAAGAGAGACACTTTG 1680  
1621 GGGTTGGGAAACAAAGAGGCATTAAGTGTATGAGCTTTGGGCAAGAGAGACACTTTG 1680  
1681 CATCTGTCTGGCAACGCTCACTCAATGCTAATGAGTGAAGCAAGCAAGCAAGCAAGTGA 1740  
1681 CATCTGTCTGGCAACGCTCACTCAATGCTAATGAGTGAAGCAAGCAAGCAAGCAAGTGA 1740  
1741 CCCCCAAGAAACCGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1741 CCCCCAAGAAACCGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1801 TCCGAGATGAAGCCCTTGCAGTATGATGAAGCAAGAGTCTTTCCCTGCTGAGGATG 1860  
1801 TCCGAGATGAAGCCCTTGCAGTATGATGAAGCAAGAGTCTTTCCCTGCTGAGGATG 1860  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTACGCTGTGGCAAGCAAGAAAGAA 1920  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTACGCTGTGGCAAGCAAGAAAGAA 1920

1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGAGA 1980  
1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGAGA 1980  
1981 CAACAGATGAGAGAGAGAGAGATCAAAATTTCCATCATGCAACCCCGGAGTCAAGTGT 2040  
1981 CAACAGATGAGAGAGAGAGAGATCAAAATTTCCATCATGCAACCCCGGAGTCAAGTGT 2040  
2041 GSCGACACAGAGGCGCCACAACTGCAGGCTCCGAATGCACCTGGGCTTGTGTATCCCA 2100  
2041 GSCGACACAGAGGCGCCACAACTGCAGGCTCCGAATGCACCTGGGCTTGTGTATCCCA 2100  
2101 AGAAGGCTGCAAGATTCGATGTGCAACAGAGACAGAGACCTGGAGAGAGCAAGGTGC 2160  
2101 AGAAGGCTGCAAGATTCGATGTGCAACAGAGACAGAGACCTGGAGAGAGCAAGGTGC 2160  
2161 TCATCTTTGATGATCTTTCGAGCAGAGGTATGCGAGATGCTCTTTCGAGCTTGA 2220  
2161 TCATCTTTGATGATCTTTCGAGCAGAGGTATGCGAGATGCTCTTTCGAGCTTGA 2220  
2221 TATTCATGCTGATGCTGCGATCCGAACTGACACCAAGAGAGAGCAAGCTTCCAG 2280  
2221 TATTCATGCTGATGCTGCGATCCGAACTGACACCAAGAGAGAGCAAGCTTCCAG 2280  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGAACTCTCGAGAGA 2324  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGAACTCTCGAGAGA 2324

RESULT 5  
US-09-859-604-3  
; Sequence 3, Application US/09859604  
; Patent No. 6835370  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M  
; APPLICANT: Deutch, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver..2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-859-604-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGGACCTGCAANTGGCCCGCAGCGTAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 60  
1 CGGACCTGCAANTGGCCCGCAGCGTAAGATGCCAAGAGCAGCGCGCAACAGCAGCAGCAGCG 60  
61 GCTCCGCGCAGCGGTAGCACCGAGTSCGGCGCAGCAGAGCCCGCGGGCCCGGAGAGAGCAA 120  
61 GCTCCGCGCAGCGGTAGCACCGAGTSCGGCGCAGCAGAGCCCGCGGGCCCGGAGAGAGCAA 120  
121 AGCATGGAGGACACAAAGATTTGGGAGAAAGCGGACCTCTCGGAACTTCTTTCACGT 180  
121 AGCATGGAGGACACAAAGATTTGGGAGAAAGCGGACCTCTCGGAACTTCTTTCACGT 180  
181 GGTATTGCTGATTTGCAATTTGCTGGCGCTCTGGAATCTGTAGCTGTGCTTTGGTTGATC 240  
181 GGTATTGCTGATTTGCAATTTGCTGGCGCTCTGGAATCTGTAGCTGTGCTTTGGTTGATC 240  
241 TTGTTGACTATGAGAGAGTTCTAGAAAACCTAGGAACTATGATGCTGATGCTGATGAG 300

Db 241 |||||TTGTTGACTATGAGGAAGTCTAGAAAACTAGGATCTATGATGCTGATGCTGATGGAG 300  
Qy 301 ATTTTGATGTGATGATGCCAAAGTTTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Db 301 ATTTTGATGTGATGATGCCAAAGTTTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCCAGAGAGCTGAGCCACACTGAGCCCGAGGAGCAGGTTCTGCTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGCTGAGCCACACTGAGCCCGAGGAGCAGGTTCTGCTGG 420  
Qy 421 AGGAGAGACCCAGAGATCGAAGATCGAAGATCGAAGATCGAAGATCGAAGATCGAAG 480  
Db 421 AGGAGAGACCCAGAGATCGAAGATCGAAGATCGAAGATCGAAGATCGAAGATCGAAG 480  
Qy 481 AAATGTTACAGCGAGAACTGTTGAGGGAGAGACTTGCACAAAGAGAGTGCACCCACAG 540  
Db 481 AAATGTTACAGCGAGAACTGTTGAGGGAGAGACTTGCACAAAGAGAGTGCACCCACAG 540  
Qy 541 GAGAACCAACAAAGAGAGATGATGAGTCTTATGCGACTGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACAAAGAGAGATGATGAGTCTTATGCGACTGATGATGATGATGATGATG 600  
Qy 601 AGACCTCGAACCTGAGATATCTCATGAAGAACCGAGCATAGTTACCGCTGGAAGAGA 660  
Db 601 AGACCTCGAACCTGAGATATCTCATGAAGAACCGAGCATAGTTACCGCTGGAAGAGA 660  
Qy 661 CAGTTTCACAGACTGTATCAGGATATGGAAGAGATGATGTTGAGCAGGAGAAATCCAG 720  
Db 661 CAGTTTCACAGACTGTATCAGGATATGGAAGAGATGATGTTGAGCAGGAGAAATCCAG 720  
Qy 721 ATTCAGTGAACCAAGTAGTAGAGATGAAGATGTGCACCATGATGATGATGATGATG 780  
Db 721 ATTCAGTGAACCAAGTAGTAGAGATGAAGATGTGCACCATGATGATGATGATGATG 780  
Qy 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840  
Qy 841 CAGAAATGAACTGCTCCCTGAGGATATCTGATGAAGATTCAGAGTAAATGATGAGAG 900  
Db 841 CAGAAATGAACTGCTCCCTGAGGATATCTGATGAAGATTCAGAGTAAATGATGAGAG 900  
Qy 901 AAGTAAGCAATTTCTGTTGAGAGAACAGCAGGAGTACCAAGAGAACAAATAGAAAA 960  
Db 901 AAGTAAGCAATTTCTGTTGAGAGAACAGCAGGAGTACCAAGAGAACAAATAGAAAA 960  
Qy 961 CAGATGATCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCTAACTTTTAAATTAAT 1020  
Db 961 CAGATGATCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCTAACTTTTAAATTAAT 1020  
Qy 1021 TTGATAAGACTATTAAAGCTGAATCTGATGCTGAGAGAAATCTCCGTAAGAGGAAAA 1080  
Db 1021 TTGATAAGACTATTAAAGCTGAATCTGATGCTGAGAGAAATCTCCGTAAGAGGAAAA 1080  
Qy 1081 TTGAGGAGCAGTGAATGCAATTTAAGAACTAGTACGCAAAATACCTCAGAGTCCAGAG 1140  
Db 1081 TTGAGGAGCAGTGAATGCAATTTAAGAACTAGTACGCAAAATACCTCAGAGTCCAGAG 1140  
Qy 1141 CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGTAAATGAGG 1200  
Db 1141 CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGTAAATGAGG 1200  
Qy 1201 TGCTAGGTGAGGCCATCCAGACCTTACCAAGAGGTGGCCAGCTTCTGATCTCCCTCAG 1260  
Db 1201 TGCTAGGTGAGGCCATCCAGACCTTACCAAGAGGTGGCCAGCTTCTGATCTCCCTCAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTGAGGCTGCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTGAGGCTGCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320  
Qy 1321 GAGTTCCCTGCTTACCTCGCAGAGATGATCAACTATTTCCCAATGATATCTTCCTTAA 1380

Db 1321 GAGGTTCCCTGCTTACCCTGCAGAGATTAGTTCAACTATTTCCCAATGATATCTTCCTTAA 1380  
Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGAACAATGCAAAAGAAAGTTT 1440  
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGAACAATGCAAAAGAAAGTTT 1440  
Qy 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTCA 1500  
Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTCA 1500  
Qy 1501 TCCTGAAGGACAGAACAAAATTTGCTGAGAGCATCCCATATTTTAAAGAGAGGATAGAT 1560  
Db 1501 TCCTGAAGGACAGAACAAAATTTGCTGAGAGCATCCCATATTTTAAAGAGAGGATAGAT 1560  
Qy 1561 CCGGAGATCCTGGCAGCTGATGATGGAGATTTTATTTTCCACCTGGGGATGCCATGCAGA 1620  
Db 1561 CCGGAGATCCTGGCAGCTGATGATGGAGATTTTATTTTCCACCTGGGGATGCCATGCAGA 1620  
Qy 1621 GGGTTGGGAAACAAAGAGGAGCATATAAGTGTGATGAGCTTGGGCAACAGAGAGGACATTTG 1680  
Db 1621 GGGTTGGGAAACAAAGAGGAGCATATAAGTGTGATGAGCTTGGGCAACAGAGAGGACATTTG 1680  
Qy 1681 CATCTGTCTGGCAACCGCTCACTCTACAAATGTGAATGAGCTGAAAGCAGACGCTTGGTGA 1740  
Db 1681 CATCTGTCTGGCAACCGCTCACTCTACAAATGTGAATGAGCTGAAAGCAGACGCTTGGTGA 1740  
Qy 1741 CCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAATCGAAAGTTAA 1800  
Db 1741 CCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAATCGAAAGTTAA 1800  
Qy 1801 TCCGAGATGAGGCTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
Db 1801 TCCGAGATGAGGCTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
Qy 1861 AAAAAGCTGAGGAGAAAGGAGCTGAGGAGCTTCAAGTTCAGCTGTGGCAGCAAGGAGAGAA 1920  
Db 1861 AAAAAGCTGAGGAGAAAGGAGCTGAGGAGCTTCAAGTTCAGCTGTGGCAGCAAGGAGAGAA 1920  
Qy 1921 ATGAAAATCCTGCAAGAGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGGAGA 1980  
Db 1921 ATGAAAATCCTGCAAGAGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGGAGA 1980  
Qy 1981 CAACAGATGCAAGAGAGAGAGAGATCAATATTTCCATCATGACCCCGGAGTCAAGTGT 2040  
Db 1981 CAACAGATGCAAGAGAGAGAGAGATCAATATTTCCATCATGACCCCGGAGTCAAGTGT 2040  
Qy 2041 GCGCGCACACAGGCGCCCAAACTGAGGCTCCGAATGCACTTGGGCTTGGTGAATCCCA 2100  
Db 2041 GCGCGCACACAGGCGCCCAAACTGAGGCTCCGAATGCACTTGGGCTTGGTGAATCCCA 2100  
Qy 2101 AGGAAGCTGCAAGATTCGATGTGCCAACAGAGACCCAGGACCTGGGAGAGAGGAGTGC 2160  
Db 2101 AGGAAGCTGCAAGATTCGATGTGCCAACAGAGACCCAGGACCTGGGAGAGAGGAGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTTGGAGCAGGATGAGGAGTGCATCTATCTTTCCGGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTTGGAGCAGGATGAGGAGTGCATCTATCTTTCCGGCTGA 2220  
Qy 2221 TATTCACTGCTGATGCTGGGATCGGAGTGCATGACCCAGAGAGAGAGAGGAGTCCAG 2280  
Db 2221 TATTCACTGCTGATGCTGGGATCGGAGTGCATGACCCAGAGAGAGAGAGGAGTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324

RESULT 6  
US-09-040-485-1  
; Sequence 1, Application US/09040485  
; Patent No. 6166176  
; GENERAL INFORMATION:  
; APPLICANT: Radosevich, James A.

```

1  TITLE OF INVENTION:  A GENE ENCODING A NOVEL MARKER FOR
2  TITLE OF INVENTION:  CANCER
3  NUMBER OF SEQUENCES:  9
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
6  STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
7  STREET:  Plaza Drive
8  CITY:  Chicago
9  STATE:  Illinois
10 COUNTRY:  USA
11 ZIP:  60611-5599
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  Patent In Release #1.0, Version #1.30
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER:  US/09/040,485
19 FILING DATE:  17-MAR-1998
20 CLASSIFICATION:  530
21 ATTORNEY/AGENT INFORMATION:
22 NAME:  Martin, Alice O.
23 REGISTRATION NUMBER:  35,601
24 REFERENCE/DOCKET NUMBER:  8998/3
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE:  312-321-4200
27 TELEFAX:  312-321-4299
28 INFORMATION FOR SEQ ID NO:  1:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH:  2442 base pairs
31 TYPE:  nucleic acid
32 STRANDEDNESS:  single
33 TOPOLOGY:  linear
34 MOLECULE TYPE:  cDNA
35 FEATURE:
36 NAME/KEY:  CDS
37 LOCATION:  70..834
38 OS-09-040-485-1

```

Query Match	35.4%	Score 823.8	DB 3	Length 2442
Best Local Similarity	99.2%	Prod. No. 6.8e-226	Indels 0	Gaps 0
Matches 828	Conservative	0	Mismatches 7	
QY	121	AGCATGGAGGACACAAGAATCGGAGGAAAGCGGACTCTCGGAACTTCATCTCTTCACGT	180	
DB	5	AGCTTGAAGGACACAAGAATCGGAGGAAAGCGGACTCTCAGGAACCTTCATCTCTTCACGT	64	
QY	181	GGTTTATGGTATTCGATTGCTGGGCGCTCGACATCTGTAGCTGTCGTTGGTTGGTTGATC	240	
DB	65	GGTTTATGGTATTCGATTGCTGGGCGCTCGACATCTGTAGCTGTCGTTGGTTGGTTGATC	124	
QY	241	TTGTTGACTATCAGGAAGTTCAGAAAACTTAGAATCTATGATGCTGATGGTGTAGTAG	300	
DB	125	TTGTTGACTATCAGGAAGTTCAGAAAACTTAGAATCTATGATGCTGATGGTGTAGTAG	184	
QY	301	ATTTTGTGATGATGATGCGCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360	
DB	195	ATTTTGTGATGATGATGCGCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	244	
QY	361	CAGCAGTCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG	420	
DB	245	CAGCAGTCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG	304	
QY	421	AGGCAGAACCCCAGAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	480	
DB	305	AGGCAGAACCCCAGAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	364	
QY	481	AAATGGTACACCGAGAACATGTTGAGGGAGAGACCTTGCAACAGAGAATGCCACCCACAG	540	
DB	365	AAATGGTACACCGAGAACATGTTGAGGGAGAGACCTTGCAACAGAGAATGCCACCCACAG	424	
QY	541	GAGAACCAACAAAGAGGATGATGAGTTCTTTATGCGCATGATGTAGATGATGATTTG	600	

Db	425	GAGAACCAACAAGAGGATGAAGTTCCTTATGGCCAGCTGATGTAGATGATGATTG	484
Qy	601	AGACCCCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACCACGTGGAAGAGA	660
Db	485	AGACCCCTGGACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACCACGTGGAAGAGA	544
Qy	661	CAGTTTCAACAGACTGTAACTAGGATATGGAAGAGATGATGTCCTGAGCAGGNAATCCAG	720
Db	545	CAGTTTCAACAGACTGTAACTAGGATATGGAAGAGATGATGTCCTGAGCAGGNAATCCAG	604
Qy	721	ATTCCAGTGAACCAAGTAGTAGAAGATGGAAGATTGCACCATGATACAGATGATGTAACAT	780
Db	605	ATTCCAGTGAACCAAGTAGTAGAAGATGGAAGATTGCACCATGATACAGATGATGTAACAT	664
Qy	781	ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAAAATCA	840
Db	665	ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAAAATCA	724
Qy	841	CAGAAGTAAGTGTCCCTCCTGAGGATAATCCTGTAGAGATTCACAGGTAATTTGTAGAAG	900
Db	725	CAGAAGTAAGTGTCCCTCCTGAGGATAATCCTGTAGAGATTCACAGGTAATTTGTAGAAG	784
Qy	901	AAGTAAGCATTTTTCTGTGGAAAGACAGCAGGAAGTACCACCAAGAAAAACAAATAG	955
Db	785	AAGTAAGCATTTTTCTGTGGAAAGACAGCAGGAAGTACCACCAAGAAAAACAAATAG	839

RESULT 7

US-09-702-705-1791

; Sequence 1791, Application US/09702705

; Patent No. 6504010

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaityanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darrick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C14

; CURRENT APPLICATION NUMBER: US/09/702,705

; CURRENT FILING DATE: 2000-10-30

; NUMBER OF SEQ ID NOS: 1833

; SOFTWARE: Fast-Seq for Windows Version 3.0

; SEQ ID NO 1791

; LENGTH: 2442

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-702-705-1791

Query Match	35.2%	Score	819;	DB	4;	Length	2442;
Best Local Similarity	98.8%;	Pos.	No. 1.6e-244;				
Matches	825;	Conservative	0;	Mismatches	10;	Indels	0;
Gaps	0;						
QY	121	AGCATGGAGGACACAAGAAATGGGAGGAGAAAGCGGACTCTCGGAACTTCATTCTTCAAGT	180				
Db	5	ACCTTGAAGGACACAGAATGGGAGGAGAAAGCGGACTCTCAGGAACATTCATTCTTCAAGT	64				
QY	181	GCTTTATGGTGATTGCATTGCTGGCGCTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC	240				
Db	65	GGTTTATGGTGATTGCATTGCTGGCGCTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC	124				
QY	241	TTGTTTGACTATCAGGAAAGTTCTAGGAAACATAGGAATCTATGATGCTGATCGTGATGGAG	300				
Db	125	TTGTTTGACTATCAGGAAAGTTCTAGGAAACATAGGAATCTATGATGCTGATCGTGATGGAG	184				
QY	301	ATTTTGATGTGATGATGCCAAAGTTTTATTAGGACTTAAAGAGAGATCTACTTCCAGAGC	360				
Db	185	ATTTTGATGTGATGATGCCAAAGTTTTATTAGGACTTAAAGAGAGATCTACTTCCAGAGC	244				

QY 361 CAGCAGTCCGCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 420  
DB 245 CAGCAGTCCGCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 304  
QY 421 AGCGAGAACCCGAGAAATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCATG 480  
DB 305 AGCGAGAACCCGAGAAATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCATG 364  
QY 481 AATGGTACGCGAGACATGTTGAGGGAGAGACTTGCACACAGAGATGGACCCACAG 540  
DB 365 AATGGTACGCGAGACATGTTGAGGGAGAGACTTGCACACAGAGATGGACCCACAG 424  
QY 541 GAGAACCAACAGAGAGATGAGTTCCTATGCGGACTGATGATGATGATGATGATG 600  
DB 425 GAGAACCAACAGAGAGATGAGTTCCTATGCGGACTGATGATGATGATGATGATG 484  
QY 601 AGACCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAG 660  
DB 485 AGACCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAG 544  
QY 661 CAGTTTACAGAGACTGTAATCAGGATATGAGAGATGATGATGATGATGATGATG 720  
DB 545 CAGTTTACAGAGACTGTAATCAGGATATGAGAGATGATGATGATGATGATGATG 604  
QY 721 ATTCCAGTGAACCTGAGTATGAGAGATGAGTTCCTATGCGGACTGATGATGATGATG 780  
DB 605 ATTCCAGTGAACCTGAGTATGAGAGATGAGTTCCTATGCGGACTGATGATGATGATG 664  
QY 781 ACCAAGTCTATGAGAAACAGCAGATATGAACTCTAGAAATGAAGGGATGAAATCA 840  
DB 665 ACCAAGTCTATGAGAAACAGCAGATATGAACTCTAGAAATGAAGGGATGAAATCA 724  
QY 841 CAGAAGTAACTGCTCCCTGAGGATAATCTCTGAGAGATTCACAGGTAATGTAGAG 900  
DB 725 CAGAAGTAACTGCTCCCTGAGGATAATCTCTGAGAGATTCACAGGTAATGTAGAG 784  
QY 901 AAGTAAGCATTTTCTGCGGAGAGAACAGCAGGAGTACCCACGAGAAACAATAG 955  
DB 785 AAGTAAGCATTTTCTGCGGAGAGAACAGCAGGAGTACCCACGAGAAACAATAG 839

## RESULT 8

US-09-736-457-1791  
; Sequence 1791, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736/457  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1791

Query Match 35.2%; Score 819; DB 4; Length 2442;  
Best Local Similarity 98.8%; Pred. No. 1.6e-224;  
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 121 AGCATGAGGAGCACAGAAATGGAGGAAAGGCGGACTCTCGGAACTTCATTCTTCACT 180  
DB 5 AGCTTGAAGGAGCACAGAAATGGAGGAAAGGCGGACTCTCGGAACTTCATTCTTCACT 64  
QY 181 GGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
DB 65 GGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 124  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGATGATGATGATG 300  
DB 125 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGATGATGATGATG 184  
QY 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
DB 185 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 244  
QY 361 CAGCAGTCCGCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 420  
DB 245 CAGCAGTCCGCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 304  
QY 421 AGCGAGAACCCGAGAAATATCGAAGATGAGCAAAAGAAACAAATTCAGTCCCTTCTCATG 480  
DB 305 AGCGAGAACCCGAGAAATATCGAAGATGAGCAAAAGAAACAAATTCAGTCCCTTCTCATG 364  
QY 481 AATGGTACACGCGAGAACATGTTGAGGAGAGAGACTTGCACACAGAGATGGACCCACAG 540  
DB 365 AATGGTACACGCGAGAACATGTTGAGGAGAGAGACTTGCACACAGAGATGGACCCACAG 424  
QY 541 GAGAACCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
DB 425 GAGAACCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 484  
QY 601 AGACCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAG 660  
DB 485 AGACCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAG 544  
QY 661 CAGTTTACAGAGACTGTAATCAGGATATGAGAGATGATGATGATGATGATGATGATG 720  
DB 545 CAGTTTACAGAGACTGTAATCAGGATATGAGAGATGATGATGATGATGATGATGATG 604  
QY 721 ATTCCAGTGAACCTGAGTATGAGAGATGAGTTCCTATGCGGACTGATGATGATGATG 780  
DB 605 ATTCCAGTGAACCTGAGTATGAGAGATGAGTTCCTATGCGGACTGATGATGATGATG 664  
QY 781 ACCAAGTCTATGAGAAACAGCAGATATGAACTCTAGAAATGAAGGGATGAAATCA 840  
DB 665 ACCAAGTCTATGAGAAACAGCAGATATGAACTCTAGAAATGAAGGGATGAAATCA 724  
QY 841 CAGAAGTAACTGCTCCCTGAGGATAATCTCTGAGAGATTCACAGGTAATGTAGAG 900  
DB 725 CAGAAGTAACTGCTCCCTGAGGATAATCTCTGAGAGATTCACAGGTAATGTAGAG 784  
QY 901 AAGTAAGCATTTTCTGCGGAGAGAACAGCAGGAGTACCCACGAGAAACAATAG 955  
DB 785 AAGTAAGCATTTTCTGCGGAGAGAACAGCAGGAGTACCCACGAGAAACAATAG 839

## RESULT 9

US-09-671-325-1791  
; Sequence 1791, Application US/09671325  
; Patent No. 6667154  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 21021.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1791
; LENGTH: 2442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-1791

Query Match 35.2%; Score 819; DB 4; Length 2442;
Best Local Similarity 98.8%; Pred. No. 1.6e-224; Indels 0; Gaps 0;
Matches 825; Conservative 0; Mismatches 10;

QY 121 AGCATGGAGGACACAAAGATGGAGGAAAGCGGACTCTCGGAACTTCACTTTCACGT 180
DB 5 AGCTTGAAGGACACAAAGATGGAGGAAAGCGGACTCTCAGGAACTTCACTTTCACGT 64
QY 181 GGTATTAGGTGATGCTGCGGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 240
DB 65 GGTATTAGGTGATGCTGCGGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 124
QY 241 TTGTTGACTATGAGGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 300
DB 185 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 244
QY 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
DB 185 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 244
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 420
DB 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 304
QY 421 AGCAGAACCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 305 AGCAGAACCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 364
QY 481 AAATGGTACACGAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 540
DB 365 AAATGGTACACGAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 424
QY 541 GAGAACCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 425 GAGAACCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 484
QY 601 AGACCCCTGGAACCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 485 AGACCCCTGGAACCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 544
QY 661 CAGTTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 545 CAGTTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 604
QY 721 ATTCCAGTGAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 605 ATTCCAGTGAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 664
QY 781 ACCAAGTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 665 ACCAAGTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 724
QY 841 CAGAGTGAACCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 725 CAGAGTGAACCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
QY 901 AAGTAAGCAATTTTCTGTTGGAAGAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGG 955
DB 785 AAGTAAGCAATTTTCTGTTGGAAGAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGG 839

RESULT 10
US-09-659-521-1
; Sequence 1, Application US/09659521
; Patent No. 6727080
; GENERAL INFORMATION:
; APPLICANT: RADOSEVICH, JAMES A.
; TITLE OF INVENTION: GENE ENCODING LABYRINTHIN, A MARKER FOR CANCER
; FILE REFERENCE: 21511/91057
; CURRENT APPLICATION NUMBER: US/09/659,521
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: PCT/US99/05365
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 09/040,485
; PRIOR FILING DATE: 1998-03-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)...(834)
US-09-659-521-1

Query Match 35.2%; Score 819; DB 4; Length 2442;
Best Local Similarity 98.8%; Pred. No. 1.6e-224; Indels 0; Gaps 0;
Matches 825; Conservative 0; Mismatches 10;

QY 121 AGCATGGAGGACACAAAGATGGAGGAAAGCGGACTCTCGGAACTTCACTTTCACGT 180
DB 5 AGCTTGAAGGACACAAAGATGGAGGAAAGCGGACTCTCAGGAACTTCACTTTCACGT 64
QY 181 GGTATTAGGTGATGCTGCGGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 240
DB 65 GGTATTAGGTGATGCTGCGGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 124
QY 241 TTGTTGACTATGAGGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 300
DB 125 TTGTTGACTATGAGGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 184
QY 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
DB 185 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 244
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 420
DB 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 304
QY 421 AGCAGAACCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 305 AGCAGAACCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 364
QY 481 AAATGGTACACGAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 540
DB 365 AAATGGTACACGAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGAGGTTCTCTG 424
QY 541 GAGAACCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 425 GAGAACCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 484
QY 601 AGACCCCTGGAACCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 485 AGACCCCTGGAACCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 544
QY 661 CAGTTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 545 CAGTTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 604
QY 721 ATTCCAGTGAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 605 ATTCCAGTGAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 664







425 AGAACCCAGAAATCGAAGATGAAGCAAAAGAAACAATTCAGTCCCTTCTCCATGAAT 484  
268 AGAACCCAGAAATCGAAGATGAAGCAAAAGAAACAATTCAGTCCCTTCTCCATGAAT 209  
485 GGTACACGACAGAAATGTTGAGGGGAGAAAGACTTGC 519  
208 GGTACACGACAGAAATGTTGAGGGGAGAAAGACTTGC 174

RESULT 15  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fis  
; US-08-232-463-14

Query Match 3.3%; Score 77.2; DB 1; Length 7218;  
Best Local Similarity 7.2%; Pred. No. 9.6e-11;  
Matches 31; Conservative 238; Mismatches 161; Indels 0; Gaps 0;

QY 759 CATGATACAGATGATGAACATCAACATCAAGTCTATCAGAGAAAGAGATATGAACTCTA 818  
DB 1490 CATCACTGTAATTACCTATCTATCAAGTAGTTAAAGAGATAGAGATTTGGTACRRR 1431  
QY 819 GAAATGAAGGATAGAAATCACAGAAATCACTCTCCCTCAGGAGTAATCTCTGTAGAA 878  
DB 1430 RRR 1371  
QY 879 GATTCAAGGTAATTGTAGAGAGATGAGATTTTCTGTGGAGAGACACAGAGATA 938

67114 TTCTTACGTGGTTTATGTTGATTCCTGGCGCTTGGACATCTGTAGTGTGCTT 67055  
231 TGGTTTCACTCTTGTGACTATGAGGAAGTTCTAGGAAA 268  
67054 TGGTTTCACTCTTGTGACTATGAGGAAGTTCTAGGTA 67017

RESULT 13  
US-09-949-016-16453/c  
; Sequence 16453, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16453  
; LENGTH: 5746  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-016-16453

Query Match 3.8%; Score 88.8; DB 4; Length 5746;  
Best Local Similarity 97.8%; Pred. No. 3.8e-14;  
Matches 90; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1069 RAAGGGGAAATTCAGGAGCAGTGAATGCAATTTAAAGAACTAGTAGGCAATACCCCTC 1128  
DB 700 AACAGGGAATTTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTAGGCAATACCCCTC 641  
QY 1129 AGAGTCCAGCAGCAAGATATGGAAGGCGCAG 1160  
DB 640 AGAGTCCAGCAGCAAGATATGGAAGGCGCAG 609

RESULT 14  
US-09-949-016-122372/c  
; Sequence 122372, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122372  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-016-122372

Query Match 3.4%; Score 79; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 4.7e-12;  
Matches 85; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

[illegible]

Search completed: March 25, 2005, 14:35:49  
Job time : 412 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 14:35:53 ; Search time 9715 Seconds  
(without alignments)  
11591.346 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466.

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 100 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	6	AX146795 Sequence
2	2324	100.0	2324	9	S83325
3	2222	95.6	2452	6	CO871564
4	1766	76.0	2449	9	HSU03109
5	1164	50.1	1970	6	CO729666
6	895	38.5	2680	9	AF289489
7	684	29.4	2280	9	AF306765
8	640	27.5	2442	6	BD234762
9	640	27.5	2442	6	AR274038
10	640	27.5	2442	6	AR277619
11	640	27.5	2442	6	AR441744
12	640	27.5	2442	6	AR532049
13	640	27.5	2442	6	AX369081
14	612	26.3	1080	9	BC015518
15	450	19.4	2761	9	BC025236
16	315	13.6	366	6	CO680735
17	297	12.8	297	6	CO731371
18	213	9.2	1608	9	BC066929
19	190	8.2	132868	2	AC021323 Homo sapi

20	190	8.2	158404	9	AC090094	AC090094 Homo sapi
21	172	7.4	855	9	AF224468	AF224468 Homo sapi
22	135	6.7	186676	9	AC067881	AC067881 Homo sapi
23	135	5.8	191	6	CO657781	CO657781 Sequence
24	130	5.6	126392	9	CR626894	CR626894 Human DNA
25	130	5.6	142265	9	AC091173	AC091173 Homo sapi
26	103	4.4	941	9	AF224469	AF224469 Homo sapi
27	103	4.4	1254	9	AF184241	AF184241 Homo sapi
28	62	2.7	2739	4	BOVASBH	M31213 Bos taurus
29	60	2.6	60	6	CO541435	CO541435 Sequence
30	59	2.5	2208	6	AX771609	AX771609 Sequence
31	56	2.5	2208	9	HSATPBR	AC03747 Human mRNA
32	56	2.4	177614	2	AC143562	AC143562 Macaca mu
33	53	2.3	1777	4	CFU38414	U38414 Canis famil
34	40	1.7	556	4	AF198966	AF198966 Oryctolag
35	35	1.5	4419	10	AF289486	AF289486 Mus muscu
36	35	1.5	6024	10	H010058511	AF289215 Mus muscu
37	35	1.5	6629	10	AF289487	AF289487 Mus muscu
38	35	1.5	168569	10	AL671970	AL671970 Mouse DNA
39	32	1.4	524	10	AF289491	AF289491 Mus muscu
40	32	1.4	1763	10	BC061098	BC061098 Mus muscu
41	32	1.4	1951	10	H010058502	AF289206 Mus muscu
42	32	1.4	2617	10	AF302653	AF302653 Mus muscu
43	32	1.4	2770	10	BC015281	BC015281 Mus muscu
44	32	1.4	2789	10	AF289488	AF289488 Mus muscu
45	32	1.4	198991	10	AL771548	AL771548 Mouse DNA
46	31	1.3	566	10	AF289492	AF289492 Mus muscu
47	31	1.3	919	10	AF221854	AF221854 Mus muscu
48	31	1.3	988	10	AF223414	AF223414 Mus muscu
49	31	1.3	1013	10	AF223413	AF223413 Mus muscu
50	31	1.3	1733	10	AF289490	AF289490 Mus muscu
51	31	1.3	2432	10	AF302655	AF302655 Mus muscu
52	31	1.3	2637	10	AF302654	AF302654 Mus muscu
53	30	1.3	158289	2	AC143465	AC143465 Macaca mu
54	29	1.2	234548	2	AC095635	AC095635 Rattus no
55	29	1.2	275547	2	AC128056	AC128056 Rattus no
56	27	1.2	27	6	CO871665	CO871665 Sequence
57	27	1.2	2013	10	H010058505	AF289209 Mus muscu
58	27	1.2	252366	2	AC096178	AC096178 Rattus no
59	26	1.1	6629	10	H010058507	AF289211 Mus muscu
60	26	1.1	12095	10	H010058510	AF289214 Mus muscu
61	24	1.0	1519	10	AF016503	AF016503 Rattus no
62	24	1.0	1860	10	MMU19607	U19607 Mus musculu
63	24	1.0	2204	10	MMU16145	U16145 Mus musculu
64	24	1.0	2221	10	MUSRLK	L35268 Mus cookii
65	24	1.0	2342	10	MUSPTKRL18	D3963 Mus musculu
66	24	1.0	110000	2	AC013623_2	Continuation (3 of
67	24	1.0	110000	2	AC013623_3	Continuation (4 of
68	24	1.0	110000	8	CR382132_00	CR382132 Yarrowia
69	24	1.0	150010	5	BX663616	BX663616 Zebrafish
70	24	1.0	192578	5	AL732455	AL732455 Zebrafish
71	24	1.0	222930	2	AC099698	AC099698 Mus muscu
72	24	1.0	224556	10	AC090126	AC090126 Mus muscu
73	24	1.0	289405	2	CO871636	CO871636 Sequence
74	23	1.0	23	6	CO871636	G94809 S209P6230RC
75	23	1.0	595	11	G94809	AF289205 Mus muscu
76	23	1.0	2483	10	H010058501	AF026465 Mus muscu
77	23	1.0	3146	10	AF026465	BC053057 Mus muscu
78	23	1.0	3198	10	BC053057	BC053057 Mus muscu
79	23	1.0	57919	2	AC017856	AC017856 Drosophil
80	23	1.0	115135	10	AC112161	AC112161 Mus Muscu
81	23	1.0	120440	5	BX000467	BX000467 Zebrafish
82	23	1.0	163122	2	AC130052	AC130052 Rattus no
83	23	1.0	192136	10	AC133508	AC133508 Mus muscu
84	23	1.0	207761	3	AC093546	AC093546 Drosophil
85	23	1.0	210493	2	AC146886	AC146886 Callithri
86	23	1.0	218311	2	CR354540	CR354540 Dario rer
87	23	1.0	220755	10	AC110235	AC110235 Mus muscu
88	23	1.0	228016	2	AC146674	AC146674 Callithri
89	23	1.0	238975	2	AC168627	AC168627 Rattus no
90	23	1.0	272016	3	AE003539	AE003539 Drosophil
91	23	1.0	273077	2	AC094884	AC094884 Rattus no
92	22	0.9	855	6	CO608337	CO608337 Sequence

93	22	0.9	1285	10	BC046342	Mus muscu
94	22	0.9	1294	10	BC024048	Mus muscu
95	22	0.9	1387	10	MMU251200	Mus muscu
96	22	0.9	1414	10	BC055431	Mus muscu
97	22	0.9	1441	10	BC064069	Mus muscu
98	22	0.9	1933	6	CQ598341	Sequence
99	22	0.9	1938	10	BC012523	Mus muscu
100	22	0.9	2003	3	DROBARH2E1	M82885 Drosophila
ALIGNMENTS						
SULT 1						
AX146795 2324 bp DNA linear PAT 31-MAY-2001						
Sequence 3 from Patent WO0135102.						
AX146795						
AX146795.1 GI:14285147						
Homo sapiens (human)						
Homo sapiens						
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
1						
Wands, J.R., de la Monte, S.M., Ince, N. and Carlson, R.I.						
Diagnosis and treatment of malignant neoplasms						
Patent: WO 0135102-A 3 17-MAY-2001;						
RHODE ISLAND HOSPITAL (US)						
Location/Qualifiers						
1..2324						
/organism="Homo sapiens"						
/mol_type="unassigned DNA"						
/db_xref="taxon:9606"						
ORIGIN						
Query Match 100.0%; Score 2324; DB 6; Length 2324;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
4	1	CGGACCGTGCAATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG	60			
5	1	CGGACCGTGCAATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG	60			
6	61	GCTCGGCGAGCGGTAGCAGCGTGGCGCAGCAGCAGCGCGCGCGCGGAGAGAGACAA	120			
7	61	GCTCGGCGAGCGGTAGCAGCGTGGCGCAGCAGCAGCGCGCGCGCGGAGAGAGACAA	120			
8	121	AGCATGGAGGACACAAGAAATGGAGAAAGCGGAGCTCTCGGGAACTTCATTCTTCACGT	180			
9	121	AGCATGGAGGACACAAGAAATGGAGAAAGCGGAGCTCTCGGGAACTTCATTCTTCACGT	180			
10	181	GGTTTATGGTATGTCATTTGCTGGCGCTCTGGACATCTGTAGCTGTGCTTGGTTTGATC	240			
11	181	GGTTTATGGTATGTCATTTGCTGGCGTCTGGACATCTGTAGCTGTGCTTGGTTTGATC	240			
12	241	TTGTTGACTATAGGAAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGAATGGAG	300			
13	241	TTGTTGACTATAGGAAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGAATGGAG	300			
14	301	ATTTTGTATGGATGATGCCAAAGTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360			
15	301	ATTTTGTATGGATGATGCCAAAGTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360			
16	361	CAGCAGTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGAGGAGCAGGTTCTCTGTGG	420			
17	361	CAGCAGTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGAGGAGCAGGTTCTCTGTGG	420			
18	421	AGGCAGAACCCAGAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG	480			
19	421	AGGCAGAACCCAGAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG	480			
20	481	AAATGGTACACGCAAGAACATGTTGAGGGAGAGAACTTGCACAAAGAGATGGACCCACAG	540			

Db	1561	CCGAGATCCTGGCAGTATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGA	1620
Qy	1621	GGGTTGGGAAACAAGAGGCATATAAGTGGTATGAGCTTGGGCACAGAGAGGACATTTG	1680
Db	1621	GGGTTGGGAAACAAGAGGCATATAAGTGGTATGAGCTTGGGCACAGAGAGGACATTTG	1680
Qy	1681	CATCTGTCTGGCAGCGCTCACTCTACAACTGAACTGAAGTCTTGAAGAACTGGAAGTAA	1740
Db	1681	CATCTGTCTGGCAGCGCTCACTCTACAACTGAACTGAAGTCTTGAAGAACTGGAAGTAA	1740
Qy	1741	CCCCAAAGAAACGGGCTACACAGCTTGAAGTCTTGAAGAACTGGAAGTAA	1800
Db	1741	CCCCAAAGAAACGGGCTACACAGCTTGAAGTCTTGAAGAACTGGAAGTAA	1800
Qy	1801	TCCGAGATGAAGGCTTTCAGTGTATGATGAAGCCAAAGGCTCTTCTCTGCTGAGGATG	1860
Db	1801	TCCGAGATGAAGGCTTTCAGTGTATGATGAAGCCAAAGGCTCTTCTCTGCTGAGGATG	1860
Qy	1861	AAAACTCAGGAAAAAGGAGCTGGAGCCAGTTCACGCTGTGGCAGCAAGGAGAA	1920
Db	1861	AAAACTCAGGAAAAAGGAGCTGGAGCCAGTTCACGCTGTGGCAGCAAGGAGAA	1920
Qy	1921	ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCTCCGCA	1980
Db	1921	ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCTCCGCA	1980
Qy	1981	CAACAGGATGCAAGAGGAGCAGATCAAAATTTCCATCATCACCCTGGGCTGAGTGT	2040
Db	1981	CAACAGGATGCAAGAGGAGCAGATCAAAATTTCCATCATCACCCTGGGCTGAGTGT	2040
Qy	2041	GGCGCACAAGGCGCCCAAACTGAGGCTCCGAGTCCGAGTCCGAGTCCGAGTCCCA	2100
Db	2041	GGCGCACAAGGCGCCCAAACTGAGGCTCCGAGTCCGAGTCCGAGTCCGAGTCCCA	2100
Qy	2101	AGGAGGCTGCAAGATTCAGTGTGCGACAGGAGCAGGACCTGGGAGGAGGAGGTC	2160
Db	2101	AGGAGGCTGCAAGATTCAGTGTGCGACAGGAGCAGGACCTGGGAGGAGGAGGTC	2160
Qy	2161	TCATCTTTGATGACTCTTTGAGCAGAGGATGAGGAGTGCCTCATCTTCCGCTGA	2220
Db	2161	TCATCTTTGATGACTCTTTGAGCAGAGGATGAGGAGTGCCTCATCTTCCGCTGA	2220
Qy	2221	TATTCATGCTGGATGTGTGGCATCCGGAATGACACACAGCAGAGCAGCTTCCAG	2280
Db	2221	TATTCATGCTGGATGTGTGGCATCCGGAATGACACACAGCAGAGCAGCTTCCAG	2280
Qy	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2340
Db	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2340
RESULT 2			
LOCUS	S83325	2324 bp	mrna
DEFINITION	aspartyl (asparaginyl) beta-hydroxylase [human, hepatoblastoma cell line HepG2, mRNA Partial, 2324 nt].		
ACCESSION	S83325		
VERSION	S83325.1	GI:1911651	
KEYWORDS	aspartyl (asparaginyl) beta-hydroxylase		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2324)		
TITLE	Lavatsiere, L., Jia, S., Nishiyama, M., de la Monte, S., Stern, A.M., Wands, J.R. and Friedman, P.A.		
JOURNAL	Overexpression of human aspartyl (asparaginyl) beta-hydroxylase in hepatocellular carcinoma and cholangiocarcinoma		
MEDLINE	J. Clin. Invest. 98 (6), 1313-1323 (1996)		
PUBMED	96420598		
REMARK	8823296		
	GenBank staff at the National Library of Medicine created this entry (NCBI gibbsq 179962) from the original journal article.		

Authors note differences between this sequence and that of P. Koriath, C. Gieffers, and J. Frey: Gene 150 (2), 395-399 (1994), GenBank U03109.

Location/Qualifiers

1. .2324

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

1. .2324

/gene="aspartyl (asparaginyl) beta-hydroxylase, HAAH"

12. .2288

/gene="aspartyl (asparaginyl) beta-hydroxylase, HAAH"

/note="HAAH"

/codon\_start=1

/product="aspartyl (asparaginyl) beta-hydroxylase"

/protein\_id="AA50779.1"

/db\_xref="GI:1911652"

/translation="MAQRNNAKSSSSSSSGSGSGTSAGSSSPGARRETKHGKNG RKGLSGTFFTFWVIALGVTSVAVMFDLVVDVVEVLKGLGIYDADGDFDVID AKVLLGKRSSEPAVPEAEHPTEPEQVPEAPQNIQDEAKEQIQSLHEMTHV AEHVEDLQSDGPTGPEQDEDFLMDVDDRTFLEPEVEHEEHSYHVEVTHV SDCNDMEEMSEGENFDSPEVPERLHDHDDVTQVVEEQAVKPEKKEGIEI TEVTAPDENPVEDSQVTEVSEVPPVEEQEVPPTNRTKTDDEPKOAKVKKDLAEKR NKFDKTKAELDAEKLARKRKIEAVNAFKELVRKYPOSPRARGYSGLLTLORLVQF RSNEVLRGAIETYQVVASLPDVPADLLKLSLRSDRQOPPLGHMRGSLTLORLVQF PNTLSKNDLVGVYLLIGDNNAKVEEVLVTNDPNDFAKHYHGFLLKAKNTARS I PNLKEGIESDGTGDRFYFHLGDMORVGNKEAYKWLGHKGHFASVWQSRSLYN VNLGAQPMWTKETGYELVKSILERNMKLIRDEGLAVMDKAKGLFLEDENLRKGD WSOFTLWQGRNENACGAKPTCTILLEKFTTIGCRGQIKYSIMHFGTHVWPHGTG TNCXELRHGLGLVPEKCKIKCANETRWEEKLVLFDDSFBEHVWQDASSPFLIFIV DWHPLEITPQRRSLPAI"

Query Match 100.0%; Score 2324; DB 9; Length 2324;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCTGCAATGGCCAGCGTAAGAATGCCAAGACAGCGGCAACAGCAGCAGCG 60

Db 1 CGGACCTGCAATGGCCAGCGTAAGAATGCCAAGACAGCGGCAACAGCAGCAGCG 60

Qy 61 GCTCCGACGCGTAGCAGCGTGGGAGGAAAGCGGAGCTCTCGGAACTTCATCTTCACT 120

Db 61 GCTCCGACGCGTAGCAGCGTGGGAGGAAAGCGGAGCTCTCGGAACTTCATCTTCACT 120

Qy 121 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGAGCTCTCGGAACTTCATCTTCACT 180

Db 121 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGAGCTCTCGGAACTTCATCTTCACT 180

Qy 181 GGTTTATGGTATGCTGCTGGGCGCTCTGGACATCTGTAGCTGTCTGTTTGTGATC 240

Db 181 GGTTTATGGTATGCTGCTGGGCGCTCTGGACATCTGTAGCTGTCTGTTTGTGATC 240

Qy 241 TTGTTGACTATGAGAAAGTTCTAGGAAACTTAGGAACTTATGATCTGATGATGATGAG 300

Db 241 TTGTTGACTATGAGAAAGTTCTAGGAAACTTAGGAACTTATGATCTGATGATGAG 300

Qy 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360

Db 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360

Qy 361 CAGCATGCTCCGCGAAGAGCTGAGCCACACTGAGCCCGAGGAGGAGGTTCTCTGTTG 420

Db 361 CAGCATGCTCCGCGAAGAGCTGAGCCACACTGAGCCCGAGGAGGAGGTTCTCTGTTG 420

Qy 421 AGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

Db 421 AGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

Qy 481 AAATGTTACACCCAGCAACATCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

Db 481 AAATGTTACACCCAGCAACATCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

541 QY GAGAACACAAAGAGGATGATGAGTTCTTATGCGGACTGATGATGATGATG 600  
541 Db GAGAACACAAAGAGGATGATGAGTTCTTATGCGGACTGATGATGATGATG 600  
601 QY AGACCCCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTACCACTGCGGAAGAGA 660  
601 Db AGACCCCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTACCACTGCGGAAGAGA 660  
661 QY CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
661 Db CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
721 QY ATTCCAGTGAACAGTATGAGAGATGGAAGATGACACCATGATACAGATGATGTAACAT 780  
721 Db ATTCCAGTGAACAGTATGAGAGATGGAAGATGACACCATGATACAGATGATGTAACAT 780  
781 QY ACCAAGTCTATGAGGAACACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
781 Db ACCAAGTCTATGAGGAACACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
841 QY CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAGAAATTCACAGGTAATTTGTAGAAG 900  
841 Db CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAGAAATTCACAGGTAATTTGTAGAAG 900  
901 QY AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAAGTACACCAAGAAACAAATAGAAAAA 960  
901 Db AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAAGTACACCAAGAAACAAATAGAAAAA 960  
961 QY CAGATGATCCAGAACAAAGAAAGTTAAAGAAAAAGAGCCTTAACTTTTAAATTAAT 1020  
961 Db CAGATGATCCAGAACAAAGAAAGTTAAAGAAAAAGAGCCTTAACTTTTAAATTAAT 1020  
1021 QY TTGATAAGACTATTAAGCTGAATCTGATGCTGAGAAAATCCGTAAGAGGGGAAAAA 1080  
1021 Db TTGATAAGACTATTAAGCTGAATCTGATGCTGAGAAAATCCGTAAGAGGGGAAAAA 1080  
1081 QY TTGAGGAAGCAGTGAATGAACTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
1081 Db TTGAGGAAGCAGTGAATGAACTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
1141 QY CAAGATATGGGAAGCGGAGTGTAGAGATGATTTGGCTGAGAAGAGGAGTAATGAGG 1200  
1141 Db CAAGATATGGGAAGCGGAGTGTAGAGATGATTTGGCTGAGAAGAGGAGTAATGAGG 1200  
1201 QY TCCTAGCTGAGCCATCGAGACCTACCAAGAGTGGCCAGCTACCTCATGCTCCCTGAG 1260  
1201 Db TCCTAGCTGAGCCATCGAGACCTACCAAGAGTGGCCAGCTACCTCATGCTCCCTGAG 1260  
1261 QY ACCTGCTGAAGCTGAGTTTGAAGCGCTCGCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320  
1261 Db ACCTGCTGAAGCTGAGTTTGAAGCGCTCGCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320  
1321 QY GAGGTTCCCTGCTTACCTCGAGAGATGATTTCAACTATTTCCCAATCATACTTCCCTTAA 1380  
1321 Db GAGGTTCCCTGCTTACCTCGAGAGATGATTTCAACTATTTCCCAATCATACTTCCCTTAA 1380  
1381 QY AAAATGACCTTCGCTGGGATACCTCTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
1381 Db AAAATGACCTTCGCTGGGATACCTCTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
1441 QY ATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTAGCTTCA 1500  
1441 Db ATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTAGCTTCA 1500  
1501 QY TCCTGAAGGCACAGAACAAATGCTGAGAGATCCCATATTTAAAGGAAGAAATAGAT 1560  
1501 Db TCCTGAAGGCACAGAACAAATGCTGAGAGATCCCATATTTAAAGGAAGAAATAGAT 1560  
1561 QY CCGGAGATCTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGCAGA 1620  
1561 Db CCGGAGATCTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGCAGA 1620  
1621 QY GGGTTGGGAACAAAGAGGCATATTAAGTGGTATGAGCTTGGGCACAGAGAGGACATTTG 1680

1621 Db GGGTTGGGAACAAAGAGGCATATTAAGTGGTATGAGCTTGGGCACAGAGAGACACTTTG 1680  
1681 QY CATCTGTCTGGCAACGCTCACTCTACAAATGTGAATGGAAGTGAAGCAGACAGCTTGGTGA 1740  
1681 Db CATCTGTCTGGCAACGCTCACTCTACAAATGTGAATGGAAGTGAAGCAGACAGCTTGGTGA 1740  
1741 QY CCCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGCTTTTAAAGAAACCTGGAAGTTAA 1800  
1741 Db CCCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGCTTTTAAAGAAACCTGGAAGTTAA 1800  
1801 QY TCCGAGATGAAGGCTTGCAGTGTGAATGAAGCCAAAGGCTCTCTTCCCTGCTGAGGATG 1860  
1801 Db TCCGAGATGAAGGCTTGCAGTGTGAATGAAGCCAAAGGCTCTCTTCCCTGCTGAGGATG 1860  
1861 QY AAAAATCTGAGGAAAGAAAGGGGACTGAGCCAGTCTCAAGCTGTGGCAGCAAGGAAGAA 1920  
1861 Db AAAAATCTGAGGAAAGAAAGGGGACTGAGCCAGTCTCAAGCTGTGGCAGCAAGGAAGAA 1920  
1921 QY ATGAAAATCCCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
1921 Db ATGAAAATCCCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
1981 QY CAACAGGATGCAAGAGAGAGAGACAGATCAATATTTCCATCATGCAACCCGGGACTCACGCTG 2040  
1981 Db CAACAGGATGCAAGAGAGAGAGACAGATCAATATTTCCATCATGCAACCCGGGACTCACGCTG 2040  
2041 QY GGGCGCACACAGGGCCCAACAACTGCAGGCTCCGAATGCACTGGGCTTGGTGTATCCCA 2100  
2041 Db GGGCGCACACAGGGCCCAACAACTGCAGGCTCCGAATGCACTGGGCTTGGTGTATCCCA 2100  
2101 QY AGAAAGGCTGCAAGATTCGATGTGCAACAGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
2101 Db AGAAAGGCTGCAAGATTCGATGTGCAACAGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
2161 QY TCATCTTGTATGACTCTTTTGGACAGAGTATGCGAGATGCGAGTGCCTCATCTTTCCGCTGA 2220  
2161 Db TCATCTTGTATGACTCTTTTGGACAGAGTATGCGAGATGCGAGTGCCTCATCTTTCCGCTGA 2220  
2221 QY TATTCATCTGTGAGTGTGTGCAATCCGAACTGACACACAGCAGAGACCGAGCTTCCAG 2280  
2221 Db TATTCATCTGTGAGTGTGTGCAATCCGAACTGACACACAGCAGAGACCGAGCTTCCAG 2280  
2281 QY CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324  
2281 Db CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324

RESULT 3  
CQ871564 LOCUS 2452 bp DNA linear PAT 27-SEP-2004  
DEFINITION Sequence 55 from Patent WO2004078783.  
CQ871564 ACCESSION  
CQ871564.1 GI:52745624  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
MURPHY, F.J., SHEEHAN, D.E., KEATING, K.E., HAYES, I.E. and SEERA, L.E.  
Enzymes involved in apoptosis  
Patent: WO 2004078783-A 55 16-SEP-2004;  
Elix Therapeutics Ltd (IE)  
FEATURES  
Location/Qualifiers  
1..2452  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 95.6%; Score 2222; DB 6; Length 2452;  
Best Local Similarity 99.9%; Pred. No. 0;





QY 2161 TCATCTTTGATGACTCTTTTGGACGACGAGGTATGGCAGGATGCTCATCTTCGGCTGA 2220  
 DB 2227 TCATCTTTGATGACTCTTTTGGACGACGAGGTATGGCAGGATGCTCATCTTCGGCTGA 2286  
 QY 2221 TATTTCATCGTGGATGTGTGGCATCGGAATCGACACACACAGCAGCAGCAGCTTCCAG 2280  
 DB 2287 TATTTCATCGTGGATGTGTGGCATCGGAATCGACACACACAGCAGCAGCAGCTTCCAG 2346  
 QY 2281 CAATTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
 DB 2347 CAATTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2390

RESULT 4  
 HSU03109 2449 bp mRNA linear PRI 30-NOV-1995  
 LOCUS Human aspartyl beta-hydroxylase mRNA, complete cds.

DEFINITION U03109  
 ACCESSION U03109.1 GI:458031  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS Koriath, F., Gieffers, C. and Frey, J.  
 TITLE Cloning and characterization of the human gene encoding aspartyl  
 beta-hydroxylase

JOURNAL Gene 150 (2), 395-399 (1994)  
 MEDLINE 95121937  
 PUBMED 7821814

REFERENCE  
 AUTHORS Koriath, F.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-NOV-1993) Koriath F., Fakultät fuer Chemie-Biochemie  
 II, Universitaet Bielefeld, Universitaetsstrasse 25, Bielefeld,  
 33615, Germany

FEATURES  
 source  
 1..2449  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="As-5"  
 /cell\_line="mg63"  
 /cell\_type="osteosarcoma"  
 /clone\_lib="MG63-ZAP"  
 1..77  
 78..2351  
 /function="hydroxylation of aspartyl and asparaginyl  
 residues"  
 /codon\_start=1  
 /product="aspartyl beta-hydroxylase"  
 /protein\_id="AA82108.1"  
 /db\_xref="GI:458032"  
 /translation="MAQRNKAWSGNSSSSGSGSTAGSSSPCARBETHKHGNG  
 RKGLSGTSFTFMVWVIALGLVSVVWFDLVYEVVLGKLGIDADGDFDVID  
 AKVLLGLKERSSEAPVPEAEPTPEEQVPEAEQNIIDEAEQIQLLHBMVH  
 AEHVREDLQEDPGTPEQEDDFLMATDVRFTLEPEVSHETESHVHEETV  
 SQDCNODMEEMSEQENPDSSEPVVEDRLHDDVTYVVEQVAVPLENEGIEI  
 TVTAPENPVEDSVIYVEVSIIPVEEQVPEPTNRKTDPEQAKVKKPKLL  
 NKFDKTKAELDAEKKRKRIEAVNAFKELVKRYQSPRARGKACQEDDLAEKR  
 RSNVEVLGAIETYQEVASLPFDPADLLLSLKRSDRQFLGHMRGSLTLQRLVQLF  
 PNDTSKLDGLGVLLIGNDNKAYVEVLSVTNDFPAKVHYGFILKAQNKIAESI  
 PVLKEGIESGDPDGTGRFVFLGLDAMQVGNKEAYKVELGKRGHGFASVWQRSLIN  
 VNLKAQCPGKETGTOLVKSLEARNKLI RDEGLVMDKAKGLFLDENLREKGDW  
 SQFTLQQGRNRNACKGPKCTLLEKEPETTGCRGOIKYSIMHPGTHVWPHGTPT  
 NCLRHLGLVLPKGGCKINCANETKEEGKVLFPDSSFEHEVWQDASSPRLIPIVD  
 VMPELTPQOORSLPAI"  
 2352..2449

5'UTR  
 CDS

3'UTR  
 ORIGIN

Query Match 76.08; Score 1766; DB 9; Length 2449;  
 Best Local Similarity 99.7%; Pred. No. 0;

Matches 2316; Conservative 0; Mismatches 5; Indels 3; Gaps 2;  
 QY 1 CGGACCGTCAATGGCCAGCGCTAGAAATGCCAAGACGAGCGGCAACAGCAGCAGCAGCG 60  
 DB 67 CGGACCGTCAATGGCCAGCGCTAGAAATGCCAAGACGAGCGGCAACAGCAGCAGCAGCG 126  
 QY 61 GGTCCGGCAGCGGTAGCACGAGTGGGGCAGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
 DB 127 GCTCCGGCAGCGGTAGCACGAGTGGGGCAGCAGCAGCGCCCGGGCCCGGAGAGACAA 186  
 QY 121 AGCATGGAGCACACAGAAATGGGAAAGGGCGGACTCTCGGGAATCTTCACTTCTCAGT 180  
 DB 187 AGCATGGAGCACACAGAAATGGGAAAGGGCGGACTCTCAGAACTTCTTCTCAGT 246  
 QY 181 GGTTCATGTGATTCATTCGTCGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
 DB 247 GGTTCATGTGATTCATTCGTCGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 306  
 QY 241 TTCTTGAATAGGAAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGGAG 300  
 DB 307 TTCTTGAATAGGAAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGGAG 366  
 QY 301 ATTTTGAATGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTTACTTCAGAGC 360  
 DB 367 ATTTTGAATGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTTACTTCAGAGC 426  
 QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGTTCTGTGG 420  
 DB 427 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGTTCTGTGG 486  
 QY 421 AGCAGAACCCCAAGATATCGAAGATGAAGCAAAATTCAGTCCCTTCTCCATG 480  
 DB 487 AGCAGAACCCCAAGATATCGAAGATGAAGCAAAATTCAGTCCCTTCTCCATG 546  
 QY 481 AATTTGGTACACGACAGAACTGTTGAGGGAGAACTTGCACAAAGAGATGGAACCCACAG 540  
 DB 547 AATTTGGTACACGACAGAACTGTTGAGGGAGAACTTGCACAAAGAGATGGAACCCACAG 606  
 QY 541 GAGAACCAACAGAGAGATGATGATTTCTTATGGGACTGATGATGATGATGATGATG 600  
 DB 607 GAGAACCAACAGAGAGATGATGATTTCTTATGGGACTGATGATGATGATGATGATG 666  
 QY 601 AGACCTTGGAACTGAACTATCATGAAGAAACCGAGCATAGTTACCACTGCGGAGAGAGA 660  
 DB 667 AGACCTTGGAACTGAACTATCATGAAGAAACCGAGCATAGTTACCACTGCGGAGAGAGA 726  
 QY 661 CAGTTTCACAAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
 DB 727 CAGTTTCACAAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 786  
 QY 721 ATTCAGTGAACAGTAGTAGAAGATGAAGATTCACCATGATGATGATGATGATGATGAT 780  
 DB 787 ATTCAGTGAACAGTAGTAGAAGATGAAGATTCACCATGATGATGATGATGATGATGAT 846  
 QY 781 ACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAATCA 840  
 DB 847 ACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAATCA 906  
 QY 841 CAGAAGTAATCTGCTCCCGCTGAGGATTAATCTCTGAGAAATTCACAGGTAATTTAGAG 900  
 DB 907 CAGAAGTAATCTGCTCCCGCTGAGGATTAATCTCTGAGAAATTCACAGGTAATTTAGAG 966  
 QY 901 AAGTAAGCAATTTCTCTGAGAAACAGCAGGAGGATTCACCAAGAAACAAATAGAAAAA 960  
 DB 967 AAGTAAGCAATTTCTCTGAGAAACAGCAGGAGGATTCACCAAGAAACAAATAGAAAAA 1026  
 QY 961 CAGATGATCCAGAAACAAAGCTTAAAGAAAAAGAGCCCTAAATCTTTTAAATAAT 1020  
 DB 1027 CAGATGATCCAGAAACAAAGCTTAAAGAAAAAGAGCCCTAAATCTTTTAAATAAT 1086  
 QY 1021 TTGATAAGCAATTTAAAGCTGAACTTGTGTCGAGAAAAACTCGTAAAAAGGGGAAAAA 1080  
 DB 1087 TTGATAAGCAATTTAAAGCTGAACTTGTGTCGAGAAAAACTCGTAAAAAGGGGAAAAA 1146

QY 1081 TTGAGGAGCAGTGAATGATTTAAAGAACTAGTACGCAATACCTCAGAGTCCAGG 1140  
Db 1147 TTGAGGAGCAGTGAATGATTTAAAGAACTAGTACGCAATACCTCAGAGTCCAGG 1206  
QY 1141 CAAGATATGGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
Db 1207 CAAGATATGGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1266  
QY 1201 TGCTACGTGGAGCCTACGAGACCTACCAAGAGTGGCCAGCTCCTGATGCTCCCTCAG 1260  
Db 1267 TGCTACGTGGAGCCTACGAGACCTACCAAGAGTGGCCAGCTCCTGATGCTCCCTCAG 1326  
QY 1261 ACCCTGCTGAAGCTGAGTTTGAAGCTGCTCAGACAGGCAACAATTTCTAGTCAATGA 1320  
Db 1327 ACCCTGCTGAAGCTGAGTTTGAAGCTGCTCAGACAGGCAACAATTTCTAGTCAATGA 1386  
QY 1321 GAGGTTCCCTGTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATATCTTCCTTAA 1380  
Db 1387 GAGGTTCCCTGTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATATCTTCCTTAA 1446  
QY 1381 AAAATGACCTTGGCTGGGATACCTCTTGATAGAGATAATGACAAATGCAAGAAAGTTT 1440  
Db 1447 AAAATGACCTTGGCTGGGATACCTCTTGATAGAGATAATGACAAATGCAAGAAAGTTT 1506  
QY 1441 ATGAAGAGTGTGAGTGTGACACTTAATGATGGCTTTGCTTAAGTCCATATGCTTCA 1500  
Db 1507 ATGAAGAGTGTGAGTGTGACACTTAATGATGGCTTTGCTTAAGTCCATATGCTTCA 1566  
QY 1501 TCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCATATTTAAAGGAAAGGAATGAAT 1560  
Db 1567 TCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCATATTTAAAGGAAAGGAATGAAT 1626  
QY 1561 CCGAGATCCTGGCACTGATGGGAGATTTATTTCCACTGGGGGATGCCATCAG 1620  
Db 1627 CCGAGATCCTGGCACTGATGGGAGATTTATTTCCACTGGGGGATGCCATCAG 1686  
QY 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCAAGAGAGGACATTTG 1680  
Db 1687 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCAAGAGAGGACATTTG 1746  
QY 1681 CATCTGCTGGCAAGCTCTACTCATCAATGTGAATGAGCTGAAAGCAAGCTTTGTGGA 1740  
Db 1747 CATCTGCTGGCAAGCTCTACTCATCAATGTGAATGAGCTGAAAGCAAGCTTT-GTGG- 1804  
QY 1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTGGAAGTTAA 1800  
Db 1805 -CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTGGAAGTTAA 1863  
QY 1801 TCCGAGATGAAGGCTTTGCAGTGTATGATGATAAAGCCAAAGGCTCTCTCTGCTGAGGATG 1860  
Db 1864 TCCGAGATGAAGGCTTTGCAGTGTATGATGATAAAGCCAAAGGCTCTCTCTGCTGAGGATG 1923  
QY 1861 AAAACCTGAGGAAAGGGGACTGGGCGAGTTCAAGCTGTGCGAGCAAGAGAGAA 1920  
Db 1924 AAAACCTGAGGAAAGGGGACTGGGCGAGTTCAAGCTGTGCGAGCAAGAGAGAA 1983  
QY 1921 ATGAATGCTGCAAGAGGCTCTTAAAGCTGTAACCTTACTAGAAAAGTTTCCCGGAGA 1980  
Db 1984 ATGAATGCTGCAAGAGGCTCTTAAAGCTGTAACCTTACTAGAAAAGTTTCCCGGAGA 2043  
QY 1981 CAACAGGATGCAAGAGGACAGATCAAAATATTCATCATGCAACCCCGGACTCAGGT 2040  
Db 2044 CAACAGGATGCAAGAGGACAGATCAAAATATTCATCATGCAACCCCGGACTCAGGT 2103  
QY 2041 GCGCGCACAGAGGCGCCACAACTGCGAGCTCCGAATGACCTGGGCTGGTGAATCCCA 2100  
Db 2104 GCGCGCACAGAGGCGCCACAACTGCGAGCTCCGAATGACCTGGGCTGGTGAATCCCA 2163  
QY 2101 AGGAGGCTGCAAGATTGATGTGCAACAGAGACCAAGGACTTGGGAGGAGCAAGGTGC 2160  
Db 2164 AGGAGGCTGCAAGATTGATGTGCAACAGAGACCAAGGACTTGGGAGGAGCAAGGTGC 2223

QY 2161 TCATCTTTGATGACTCTCTTTGACACAGAGGTATGCGAGATGCTCATCTTTCCGGCTGA 2220  
Db 2224 TCATCTTTGATGACTCTCTTTGAGCAAGAGTATGGAGATGCTCATCTTTCCGGCTGA 2283  
QY 2221 TATTTCATGCTGATGTTGTGGCATCCGGAATCTGACACCAAGAGAGAGGAGCTTCCAG 2280  
Db 2284 TATTTCATGCTGATGTTGTGGCATCCGGAATCTGACACCAAGAGAGAGGAGCTTCCAG 2343  
QY 2281 CAATTTAGCATGAATTCATCAAGCTTGGGAACTCTGGAGAGA 2324  
Db 2344 CAATTTAGCATGAATTCATCAAGCTTGGGAACTCTGGAGAGA 2387

RESULT 5  
QY729666  
LOCUS QY729666 1970 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 15600 from Patent WO02068579.  
ACCESSION QY729666  
VERSION QY729666.1 GI:42301377  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof  
JOURNAL Patent: WO 02068579-A 15600 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source  
1:1970  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 50.1%; Score 1164; DB 6; Length 1970;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1214; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1074 GGAAAAATTCAGGAAGCAGTGAATGCATTTAAAGAACTAGTACGCAATACCTCAGAGT 1133  
Db 756 GGAAAAATTCAGGAAGCAGTGAATGCATTTAAAGAACTAGTACGCAATACCTCAGAGT 815  
QY 1134 CCAGCAGAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAGT 1193  
Db 816 CCAGCAGAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAGT 875  
QY 1194 AATGAGTGTCTAGTGGAGCCATCGAGACCTTACCAAGAGGTGGCCAGCCTACCTGATGTC 1253  
Db 876 AATGAGTGTCTAGTGGAGCCATCGAGACCTTACCAAGAGGTGGCCAGCCTACCTGATGTC 935  
QY 1254 CCTGCGAGCCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGT 1313  
Db 936 CCTGCGAGCCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGT 995  
QY 1314 CATATGAGAGGTTCCCTGCTTACCTCCAGAGATTAGTTCAACTATTTCCCAATGATCT 1373  
Db 996 CATATGAGAGGTTCCCTGCTTACCTCCAGAGATTAGTTCAACTATTTCCCAATGATCT 1055  
QY 1374 TCCTTTAAAAATGACCTTTGGCTGGGATACCTCTTGATAGGAGATAATGACAAATGCAAG 1433  
Db 1056 TCCTTTAAAAATGACCTTTGGCTGGGATACCTCTTGATAGGAGATAATGACAAATGCAAG 1115  
QY 1434 AAAGTTTATGAAGAGGCTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTAT 1493  
Db 1116 AAAGTTTATGAAGAGGCTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTAT 1175  
QY 1494 GGCTTCATCTCAAGGCGACAGCAAAATTTGCTGAGAGCATCCCATATTTTAAAGGAGGA 1553  
Db 1176 GGCTTCATCTCAAGGCGACAGCAAAATTTGCTGAGAGCATCCCATATTTTAAAGGAGGA 1235



Db 481 AATGGTACACGAGCAACATGTTGAGGAGAGAGACTTGGCAAGAAGATGACCCACAG 540  
 QY 541 GAGAACCAACAAGAGGATGATGAGTTCTTATGGCGACTGATGATGATGATGATGATG 600  
 Db 541 GAGAACCAACAAGAGGATGATGAGTTCTTATGGCGACTGATGATGATGATGATGATG 600  
 QY 601 AGACCTCGAAGCTGAGTATCTCATGAAGAACCGGAGATGATGATGATGATGATGATG 660  
 Db 601 AGACCTCGAAGCTGAGTATCTCATGAAGAACCGGAGATGATGATGATGATGATGATG 660  
 QY 661 CAGTTTCAACAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 Db 661 CAGTTTCAACAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 QY 721 ATTCCAGTGAACAGCAGTAGTAGAAGATGAAAGATGGAAGATGGAAGATGGAAGATG 780  
 Db 721 ATTCCAGTGAACAGCAGTAGTAGAAGATGGAAGATGGAAGATGGAAGATGGAAGATG 780  
 QY 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTGAGAAATGAGGAGATGAATCA 840  
 Db 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTGAGAAATGAGGAGATGAATCA 840  
 QY 841 CAGAAGTAAGTCTCTCCCTCAGGATTAATCTGTAGAGATTAATCTGTAGAGATTAATCT 900  
 Db 841 CAGAAGTAAGTCTCTCCCTCAGGATTAATCTGTAGAGATTAATCTGTAGAGATTAATCT 900  
 QY 901 AGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGATGACCAACAG 946  
 Db 901 AGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGATGACCAACAG 946

## RESULT 7

AF306765 2280 bp mRNA linear PRI 24-DEC-2000  
 LOCUS Homo sapiens junctate mRNA, complete cds.  
 DEFINITION AF306765  
 ACCESSION AF306765.1 GI:11991236  
 VERSION  
 KEYWORDS  
 SOURCE

## ORGANISM

Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2280)  
 Treves,S., Feriotto,G., Moccagatta,L., Gambari,R. and Zorzato,F.  
 Molecular cloning, expression, functional characterization, a novel  
 chromosomal localization, and gene structure of junctate, a novel  
 integral calcium binding protein of sarco(endo)plasmic reticulum  
 membrane  
 J. Biol. Chem. 275 (50), 39555-39568 (2000)  
 MEDLINE 20564329  
 PUBMED 11007777

## REFERENCE

2 (bases 1 to 2280)  
 Treves,S., Feriotto,G., Moccagatta,L., Gambari,R. and Zorzato,F.  
 Direct Submission  
 Submitted (19-SEP-2000) Experimental and Anesthesia and Research,  
 University of Basel, Hebelstrasse 22, Basel 4031, Switzerland  
 Location/Qualifiers

## FEATURES

## Source

1..2280  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8q12.1"  
 83..982  
 /note="expressed cardiac and skeletal muscle"  
 /codon\_start=1  
 /product="junctate"  
 /protein\_id="AA042257.1"  
 /db\_xref="GI:11991237"  
 /translation="MAEDKETHKHGKNGKGLSGTSFPTFWVIALGLVYTSVAVV  
 WFLDYEVFLAKDFRYNLSEVLQKGLGYDADGDGDFVDVDAKVLGKERSTSE  
 PAVPPEAPHPTEPEQVPEAEQNTDEAKQIQSLIHENVHAHVHGKEDLQEDG  
 PTGEPOQEDFLMATDVRDFTELEFVSHETSHVETVSDQNCQNMEEWSE

## CDS

QENPDSSEPVVVEERLHDTDDVTYQVVEQAVYEPLENEGIEITEVTAPPEDNPVED  
 SOVIVEESIFPVVEEQVEPPDT"

## ORIGIN

Query Match 29.4%; Score 684; DB 9; Length 2280;  
 Best local Similarity 100.0%; Pred. No. 0;  
 Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 263 AGGAAACCTAGGAATCTATGATCTGATGATGATGATGATGATGATGATGATGATG 322  
 Db 292 AGGAAACCTAGGAATCTATGATCTGATGATGATGATGATGATGATGATGATGATG 351  
 QY 323 AGTTTATATAGGACTTAAAGAGAGATCTACTTCAGAGCCAGGATGTCCTGGCCAGAGGC 382  
 Db 352 AGTTTATATAGGACTTAAAGAGAGATCTACTTCAGAGCCAGGATGTCCTGGCCAGAGGC 411  
 QY 383 TGAGCCACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCGAAGACCCAGAAATATCGA 442  
 Db 412 TGAGCCACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCGAAGACCCAGAAATATCGA 471  
 QY 443 AGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAAATGGTACACGCAACATGT 502  
 Db 472 AGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAAATGGTACACGCAACATGT 531  
 QY 503 TGAGGGAGAGACTTGCACAGAGAGATGGACCCAGAGAGAACCAACAAAGAGATGA 562  
 Db 532 TGAGGGAGAGACTTGCACAGAGAGATGGACCCAGAGAGAACCAACAAAGAGATGA 591  
 QY 563 TGAGTTTCTTATGGCGACTGTGTAGATGATGATGATGATGATGATGATGATGATGATG 622  
 Db 592 TGAGTTTCTTATGGCGACTGTGTAGATGATGATGATGATGATGATGATGATGATGATG 651  
 QY 623 TCATGAAGAAACCGAGCATAGTTTACACGCTGGAAGAGACAGATTTTCAAGAGCTGTAAATCA 682  
 Db 652 TCATGAAGAAACCGAGCATAGTTTACACGCTGGAAGAGACAGATTTTCAAGAGCTGTAAATCA 711  
 QY 683 GGATATGGAAGAGATGATGCTGAGCAGGAAATCCAGATTCAGATGATGATGATGATGATG 742  
 Db 712 GGATATGGAAGAGATGATGCTGAGCAGGAAATCCAGATTCAGATGATGATGATGATGATG 771  
 QY 743 AGATGAAGATTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 802  
 Db 772 AGATGAAGATTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 831  
 QY 803 AGTATATGAACCTCTAGAAAATGAAGGATAGAAAATCACAGAAATGAACTGCTCCCTCTGA 862  
 Db 832 AGTATATGAACCTCTAGAAAATGAAGGATAGAAAATCACAGAAATGAACTGCTCCCTCTGA 891  
 QY 863 GGATAATCTCTGTAGAGATTCACAGGTAATTTGTAGAAAGATGAAGCATTTTTCCTGTGGA 922  
 Db 892 GGATAATCTCTGTAGAGATTCACAGGTAATTTGTAGAAAGATGAAGCATTTTTCCTGTGGA 951  
 QY 923 AGAACAGCAGGAAGTACCAACACAGA 946  
 Db 952 AGAACAGCAGGAAGTACCAACACAGA 975

## RESULT 8

BD234762 2442 bp DNA linear PAT 17-JUL-2003  
 LOCUS Gene encoding cancer marker labyrinthine.  
 DEFINITION BD234762  
 ACCESSION BD234762  
 VERSION BD234762.1 GI:33044532  
 KEYWORDS JP 2002512005-A/1.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

## REFERENCE

1 (bases 1 to 2442)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Radosevich,J.A.  
 Gene encoding cancer marker labyrinthine  
 Patent: JP 2002512005-A 1 23-APR-2002;  
 JAMES A RADOSEVICH

```
COMMENT OS Homo sapiens (human)
PN JP 2002512005-A/1
PF 23-APR-2002
PR 11-MAR-1999 JP 2000536866
PR 17-MAR-1998 US 09/040485
PI JAMES A RADOSEVICH
PC C12N15/09,C07K14/47,C07K16/18,C12P21/08,C12Q1/02,C12Q1/68// PC
CC (C12P21/08,C12R1/91),C12N15/00
FH Gene encoding cancer marker labyrinthine
FT Key Location/Qualifiers
CDS CDS (70)..(834).

FEATURES             Location/Qualifiers
     source            1..2442
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"

ORIGIN
Query Match       27.5%; Score 640; DB 6; Length 2442;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAACCTTCATTCTTCACGCTGGTTTATGCTGATTCGATTCGCTGGCGCTCTGGACATCTGTA 221
Db 46 GGAACCTTCATTCTTCACGCTGGTTTATGCTGATTCGATTCGCTGGCGCTCTGGACATCTGTA 105

QY 222 GCTGCTGTTTGGTTTGATCTTGACATGAGGAGTTCTAGGAAACTAGGAATCTAT 281
Db 106 GCTGCTGTTTGGTTTGATCTTGACATGAGGAGTTCTAGGAAACTAGGAATCTAT 165

QY 282 GATGCTGATGGTGATGGAGATTTTCATGTGGATGATGCCAAAGTTTATTAGGACTTAAA 341
Db 166 GATGCTGATGGTGATGGAGATTTTCATGTGGATGATGCCAAAGTTTATTAGGACTTAAA 225

QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 401
Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 285

QY 402 GAGGAGCAGGTTCTGTGGAGCGAGACCCAGAGATATCGAAGATGAAGCAAAAGAACAA 461
Db 286 GAGGAGCAGGTTCTGTGGAGCGAGACCCAGAGATATCGAAGATGAAGCAAAAGAACAA 345

QY 462 ATTCACTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGGGAGAGACTTGCAA 521
Db 346 ATTCACTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGGGAGAGACTTGCAA 405

QY 522 CAAGAAGATGGACCCACAGAGAACCCAGAGGATGATGAGTTTCTTATGGCGACT 581
Db 406 CAAGAAGATGGACCCACAGAGAACCCAGAGGATGATGAGTTTCTTATGGCGACT 465

QY 582 GATGTAGATGATAGATTTTCAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT 641
Db 466 GATGTAGATGATAGATTTTCAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT 525

QY 642 AGTTACCACTGGAAGAGACAGTTTCAAGACTGATAGAGATGAAGAGATGATG 701
Db 526 AGTTACCACTGGAAGAGACAGTTTCAAGACTGATAGAGATGAAGAGATGATG 585

QY 702 TCTGAGCAGGAAATCCAGATTCAGTGACCACTAGTAGAGATGAAGAGATGGACCAT 761
Db 586 TCTGAGCAGGAAATCCAGATTCAGTGACCACTAGTAGAGATGAAGAGATGGACCAT 645

QY 762 GATACAGATGATGTAACATACCAAGTCTATGAGGAACAAAGCAGTATATGAACCTCTAGAA 821
Db 646 GATACAGATGATGTAACATACCAAGTCTATGAGGAACAAAGCAGTATATGAACCTCTAGAA 705

QY 822 AATGAAGGGATAGAAATCACAGAAGTAACTG 852
Db 706 AATGAAGGGATAGAAATCACAGAAGTAACTG 736
```

```
LOCUS AR274038 2442 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1791 from patent US 6504010.
ACCESSION AR274038
VERSION AR274038.1 GI:29705923
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2442)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,M.W., Mannion,J. and Fan,L.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6504010-A 1791 07-JAN-2003;
FEATURES             Location/Qualifiers
     source            1..2442
                        /organism="unknown"
                        /mol_type="genomic DNA"

ORIGIN
Query Match       27.5%; Score 640; DB 6; Length 2442;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAACCTTCATTCTTCACGCTGGTTTATGCTGATTCGATTCGCTGGCGCTCTGGACATCTGTA 221
Db 46 GGAACCTTCATTCTTCACGCTGGTTTATGCTGATTCGATTCGCTGGCGCTCTGGACATCTGTA 105

QY 222 GCTGCTGTTTGGTTTGATCTTGACATGAGGAGTTCTAGGAAACTAGGAATCTAT 281
Db 106 GCTGCTGTTTGGTTTGATCTTGACATGAGGAGTTCTAGGAAACTAGGAATCTAT 165

QY 282 GATGCTGATGGTGATGGAGATTTTCATGTGGATGATGCCAAAGTTTATTAGGACTTAAA 341
Db 166 GATGCTGATGGTGATGGAGATTTTCATGTGGATGATGCCAAAGTTTATTAGGACTTAAA 225

QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 401
Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 285

QY 402 GAGGAGCAGGTTCTGTGGAGCGAGACCCAGAGATATCGAAGATGAAGCAAAAGAACAA 461
Db 286 GAGGAGCAGGTTCTGTGGAGCGAGACCCAGAGATATCGAAGATGAAGCAAAAGAACAA 345

QY 462 ATTCACTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGGGAGAGACTTGCAA 521
Db 346 ATTCACTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGGGAGAGACTTGCAA 405

QY 522 CAAGAAGATGGACCCACAGAGAACCCAGAGGATGATGAGTTTCTTATGGCGACT 581
Db 406 CAAGAAGATGGACCCACAGAGAACCCAGAGGATGATGAGTTTCTTATGGCGACT 465

QY 582 GATGTAGATGATAGATTTTCAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT 641
Db 466 GATGTAGATGATAGATTTTCAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT 525

QY 642 AGTTACCACTGGAAGAGACAGTTTCAAGACTGATAGAGATGAAGAGATGATG 701
Db 526 AGTTACCACTGGAAGAGACAGTTTCAAGACTGATAGAGATGAAGAGATGATG 585

QY 702 TCTGAGCAGGAAATCCAGATTCAGTGACCACTAGTAGAGATGAAGAGATGGACCAT 761
Db 586 TCTGAGCAGGAAATCCAGATTCAGTGACCACTAGTAGAGATGAAGAGATGGACCAT 645

QY 762 GATACAGATGATGTAACATACCAAGTCTATGAGGAACAAAGCAGTATATGAACCTCTAGAA 821
Db 646 GATACAGATGATGTAACATACCAAGTCTATGAGGAACAAAGCAGTATATGAACCTCTAGAA 705

QY 822 AATGAAGGGATAGAAATCACAGAAGTAACTG 852
Db 706 AATGAAGGGATAGAAATCACAGAAGTAACTG 736
```

## RESULT 10

AR277619  
LOCUS  
DEFINITION  
AR277619 Sequence 1791 from patent US 6509448.  
ACCESSION  
AR277619.1 GI:29711268  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
Unclassified.  
REFERENCE  
1 (bases 1 to 2442)  
AUTHORS  
Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,  
Carter, D., Retter, M.W., Mannion, J., Fan, L.,  
TITLE  
Compositions and methods for the therapy and diagnosis of lung  
cancer  
JOURNAL  
Patent: US 6509448-A 1791 21-JAN-2003;  
FEATURES  
Location/Qualifiers  
1..2442  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 27.5%; Score 640; DB 6; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 162 GGAATTCATTCTTACGTCGTTTATGGTATGTCATTGCTGGCGCTCGGACATCTGTA 221  
Db 46 GGAATTCATTCTTACGTCGTTTATGGTATGTCATTGCTGGCGCTCGGACATCTGTA 105  
QY 222 GCTGTCGTTTGGTTTGTATCTTGTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTAT 281  
Db 106 GCTGTCGTTTGGTTTGTATCTTGTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTAT 165  
QY 282 GATGTCGTGATGATGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 166 GATGTCGTGATGATGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 225  
QY 342 GAGAGATCTACTTTCAGAGCCAGCAGTCCCGCAGAGAGAGCTGAGCCACACACTGAGCCC 401  
Db 226 GAGAGATCTACTTTCAGAGCCAGCAGTCCCGCAGAGAGAGCTGAGCCACACACTGAGCCC 285  
QY 402 GAGGAGCAGGTTCTCTGAGGAGCAGAACCCCGCAGAGATATCGAAGATGAAGCAAGAACAA 461  
Db 286 GAGGAGCAGGTTCTCTGAGGAGCAGAACCCCGCAGAGATATCGAAGATGAAGCAAGAACAA 345  
QY 462 ATTCAGTCCCTTCTCCATGAAATGGTACACGAGAACCATGTTGAGGAGAGAGACTTGCAG 521  
Db 346 ATTCAGTCCCTTCTCCATGAAATGGTACACGAGAACCATGTTGAGGAGAGAGACTTGCAG 405  
QY 522 CAAGAAGATGGAACCCACAGGAGAACCAACAAAGAGAGATGATGATGATGATGATGATGATGAT 581  
Db 406 CAAGAAGATGGAACCCACAGGAGAACCAACAAAGAGAGATGATGATGATGATGATGATGATGAT 465  
QY 582 GAT 641  
Db 466 GAT 525  
QY 642 AGTTACCACTGGAAGAGACAGTTTTCACAGACTCTGTAATCAGGATATGGAAGAGATGATG 701  
Db 526 AGTTACCACTGGAAGAGACAGTTTTCACAGACTCTGTAATCAGGATATGGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAGAAATCCAGATTCAGTGAAACAGTAGTAGGATGAAAGATGGAACCAT 761  
Db 586 TCTGAGCAGGAGAAATCCAGATTCAGTGAAACAGTAGTAGGATGAAAGATGGAACCAT 645  
QY 762 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 821  
Db 646 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 705  
QY 822 AATGAAGGGATAGAAATCACAAGAGTAACTG 852  
Db 706 AATGAAGGGATAGAAATCACAAGAGTAACTG 736

## RESULT 11

AR441744  
LOCUS  
DEFINITION  
AR441744 Sequence 1791 from patent US 6667154.  
ACCESSION  
AR441744.1 GI:42667904  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
Unclassified.  
REFERENCE  
1 (bases 1 to 2442)  
AUTHORS  
Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,  
Carter, D., Retter, M.W., Mannion, J., and Fan, L.  
TITLE  
Compositions and methods for the therapy and diagnosis of lung  
cancer  
JOURNAL  
Patent: US 6667154-A 1791 23-DEC-2003;  
FEATURES  
Location/Qualifiers  
1..2442  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 27.5%; Score 640; DB 6; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 162 GGAATTCATTCTTACGTCGTTTATGGTATGTCATTGCTGGCGCTCGGACATCTGTA 221  
Db 46 GGAATTCATTCTTACGTCGTTTATGGTATGTCATTGCTGGCGCTCGGACATCTGTA 105  
QY 222 GCTGTCGTTTGGTTTGTATCTTGTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTAT 281  
Db 106 GCTGTCGTTTGGTTTGTATCTTGTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTAT 165  
QY 282 GATGTCGTGATGATGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 166 GATGTCGTGATGATGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 225  
QY 342 GAGAGATCTACTTTCAGAGCCAGCAGTCCCGCAGAGAGAGCTGAGCCACACACTGAGCCC 401  
Db 226 GAGAGATCTACTTTCAGAGCCAGCAGTCCCGCAGAGAGAGCTGAGCCACACACTGAGCCC 285  
QY 402 GAGGAGCAGGTTCTCTGAGGAGCAGAACCCCGCAGAGATATCGAAGATGAAGCAAGAACAA 461  
Db 286 GAGGAGCAGGTTCTCTGAGGAGCAGAACCCCGCAGAGATATCGAAGATGAAGCAAGAACAA 345  
QY 462 ATTCAGTCCCTTCTCCATGAAATGGTACACGAGAACCATGTTGAGGAGAGAGACTTGCAG 521  
Db 346 ATTCAGTCCCTTCTCCATGAAATGGTACACGAGAACCATGTTGAGGAGAGAGACTTGCAG 405  
QY 522 CAAGAAGATGGAACCCACAGGAGAACCAACAAAGAGAGATGATGATGATGATGATGATGATGAT 581  
Db 406 CAAGAAGATGGAACCCACAGGAGAACCAACAAAGAGAGATGATGATGATGATGATGATGATGAT 465  
QY 582 GAT 641  
Db 466 GAT 525  
QY 642 AGTTACCACTGGAAGAGACAGTTTTCACAGACTCTGTAATCAGGATATGGAAGAGATGATG 701  
Db 526 AGTTACCACTGGAAGAGACAGTTTTCACAGACTCTGTAATCAGGATATGGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAGAAATCCAGATTCAGTGAAACAGTAGTAGGATGAAAGATGGAACCAT 761  
Db 586 TCTGAGCAGGAGAAATCCAGATTCAGTGAAACAGTAGTAGGATGAAAGATGGAACCAT 645  
QY 762 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 821  
Db 646 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 705  
QY 822 AATGAAGGGATAGAAATCACAAGAGTAACTG 852

```
|||||
706 AATGAAGGATAGAAATCACAGAACTAAGTAACTG 736

Db

RESULT 12
AR532049
LOCUS AR532049 2442 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6727080.
ACCESSION AR532049
VERSION AR532049.1 GI:53920638
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2442)
AUTHORS Radosevich, J.
TITLE Gene encoding labyrinthin, a marker for cancer
JOURNAL Patent: US 6727080-A 1 27-APR-2004;
FEATURES Location/Qualifiers
source 1..2442
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 27.5%; Score 640; DB 6; Length 2442;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCCAGTGGTTTATGGTATGATTCGATTCCTGGGGCTCTGGACATCTGTA 221
Db 46 GGAATTCATCTTCCAGTGGTTTATGGTATGATTCGATTCCTGGGGCTCTGGACATCTGTA 105
QY 222 GCTGTCGTTTGGTTTGATCTTGTTCACATAGGAAAGTTCTAGGAAAACTAGGAATCTAT 281
Db 106 GCTGTCGTTTGGTTTGATCTTGTTCACATAGGAAAGTTCTAGGAAAACTAGGAATCTAT 165
QY 282 GATGCTGATGTTGATGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 341
Db 166 GATGCTGATGTTGATGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 225
QY 342 GAGAGATCTACTTCAGAGCCAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401
Db 226 GAGAGATCTACTTCAGAGCCAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 285
QY 402 GAGGAGCAGGTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 461
Db 286 GAGGAGCAGGTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 345
QY 462 ATTCAGTCCCTTCTCCATGAAATGGTACACGAGAAACATGTTGAGGGAGAGAGCTTGCAA 521
Db 346 ATTCAGTCCCTTCTCCATGAAATGGTACACGAGAAACATGTTGAGGGAGAGAGCTTGCAA 405
QY 522 CAAGAAGATGGACCCACAGGAGAACCAACAGAGAGGATGATGATGATGATGATGATGATGATG 581
Db 406 CAAGAAGATGGACCCACAGGAGAACCAACAGAGAGGATGATGATGATGATGATGATGATG 465
QY 582 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 641
Db 466 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 525
QY 642 AGTTACACAGTGGAGAGACAGATTTTCAAGACTGTAATCAGGATATGGAAGAGATGATG 701
Db 526 AGTTACACAGTGGAGAGACAGATTTTCAAGACTGTAATCAGGATATGGAAGAGATGATG 585
QY 702 TCTGAGCAGGAAAAATCCAGATTCAGTGAACCCAGTAGTAGAGATGGAAGATTTGCACCAT 761
Db 586 TCTGAGCAGGAAAAATCCAGATTCAGTGAACCCAGTAGTAGAGATGGAAGATTTGCACCAT 645
QY 762 GATACAGATGATGTAACATACCAAGTCTATAGGAAACAGCAGTATATGAACCTCTAGAA 821
Db 646 GATACAGATGATGTAACATACCAAGTCTATAGGAAACAGCAGTATATGAACCTCTAGAA 705
QY 822 AATGAAGGATAGAAATCACAGAACTAAGTAACTG 852
```

```
|||||
706 AATGAAGGATAGAAATCACAGAACTAAGTAACTG 736

Db

RESULT 13
AX369081
LOCUS AX369081 2442 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 1791 from Patent WO0204514.
ACCESSION AX369081
VERSION AX369081.1 GI:18857144
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W., Marnerakis, M., Carter, D., Fanger, G.R., Vedwick, T.S., Bangur, C.S., McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0204514-A 1791 17-JAN-2002;
FEATURES Location/Qualifiers
source 1..2442
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 27.5%; Score 640; DB 6; Length 2442;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCCAGTGGTTTATGGTATGATTCGATTCCTGGGGCTCTGGACATCTGTA 221
Db 46 GGAATTCATCTTCCAGTGGTTTATGGTATGATTCGATTCCTGGGGCTCTGGACATCTGTA 105
QY 222 GCTGTCGTTTGGTTTGATCTTGTTCACATAGGAAAGTTCTAGGAAAACTAGGAATCTAT 281
Db 106 GCTGTCGTTTGGTTTGATCTTGTTCACATAGGAAAGTTCTAGGAAAACTAGGAATCTAT 165
QY 282 GATGCTGATGTTGATGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 341
Db 166 GATGCTGATGTTGATGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 225
QY 342 GAGAGATCTACTTCAGAGCCAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401
Db 226 GAGAGATCTACTTCAGAGCCAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 285
QY 402 GAGGAGCAGGTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 461
Db 286 GAGGAGCAGGTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 345
QY 462 ATTGAGTCCCTTCTCCATGAAATGGTACACGAGAAACATGTTGAGGGAGAGAGCTTGCAA 521
Db 346 ATTGAGTCCCTTCTCCATGAAATGGTACACGAGAAACATGTTGAGGGAGAGAGCTTGCAA 405
QY 522 CAAGAAGATGGACCCACAGGAGAACCAACAGAGAGGATGATGATGATGATGATGATGATGATG 581
Db 406 CAAGAAGATGGACCCACAGGAGAACCAACAGAGAGGATGATGATGATGATGATGATGATG 465
QY 582 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 641
Db 466 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 525
QY 642 AGTTACACAGTGGAGAGACAGATTTTCAAGACTGTAATCAGGATATGGAAGAGATGATG 701
Db 526 AGTTACACAGTGGAGAGACAGATTTTCAAGACTGTAATCAGGATATGGAAGAGATGATG 585
QY 702 TCTGAGCAGGAAAAATCCAGATTCAGTGAACCCAGTAGTAGAGATGGAAGATTTGCACCAT 761
Db 586 TCTGAGCAGGAAAAATCCAGATTCAGTGAACCCAGTAGTAGAGATGGAAGATTTGCACCAT 645
```







Db 718 GAGAACCAACAAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATG 777  
Qy 601 AGACCTCGAACCTCAAGTATCTCATGAAGAAACCGACGATAGTTACCACTGGAAGAGA 660  
Db 778 AGACCTCGAACCTCAAGTATCTCATGAAGAAACCGACGATAGTTACCACTGGAAGAGA 837  
Qy 661 CAG 663  
Db 838 CAG 840

RESULT 15  
BC025236  
LOCUS  
DEFINITION Homo sapiens aspartate beta-hydroxylase, mRNA (cdna clone  
IMAGE:5090911), complete cds.  
ACCESSION BC025236  
VERSION BC025236.1 GI:19263910  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2761)  
Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Wax,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavani,T.L.,  
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Rulyk,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skaleka,U., Smailus,D.E.,  
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932  
2 (bases 1 to 2761)  
Strausberg,R.  
Direct Submission  
Submitted (05-MAR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ruben Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Liisa Parbhui, Parvaneh Saeedi, Jacqueline  
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 42 Row: n Column: 19  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 14589865  
This clone has the following problem: The cds is short compared to  
the longest cds in the locus.  
FEATURES  
Location/Qualifiers  
1..2761  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5090911"  
/tissue\_type="Pancreas, epithelioid carcinoma"  
/clone\_lib="NIH\_MGC\_42"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1..2761  
/gene="ASPH"  
/note="synonyms: JCTN, HAAH, CASQ2BP1, BAH"  
/db\_xref="LocusID:444"  
198..1010  
/db\_xref="MIM:600582"  
/gene="ASPH"  
/codon\_start=1  
/product="ASPH protein"  
/protein\_id="AAH25236.1"  
/db\_xref="GI:19263911"  
/db\_xref="LocusID:444"  
/db\_xref="MIM:600582"  
/translation="MAQRKAKSGNSSGSGSTAGSAGSSPGARRETKHGKHNG  
RKGLSGTSFTFMVIALGVMTSAVVMFDLYDEVVLGKLIYDADGGDFDMDH  
AKVLLGKSTSEPAVPEEAPHTPEFQVPEAPQIEDEAKQIQSLLEHVDH  
AEHTEHSYHVEETVSQDCNQDMEMSEQENPDSFVVEDELRHDTDDVTQVYE  
EQAVYEPLNEGIEITEVTAPPEDNPVEDSQVIVEVSIFFVEEQSEVPEDT"  
ORIGIN  
Query Match 19.4%; Score 450; DB 9; Length 2761;  
Best Local Similarity 99.8%; Pred. No. 1.3e-235;  
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCAATGCCCGCAGCGTAAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 60  
Db 187 CGGACCGTGCAATGCCCGCAGCGTAAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 246  
Qy 61 GCTCCGCGAGCGGTAGCACCGAGTCGGGCGAGCAGCAGCCCCCGGCGCGGAGAGACAA 120  
Db 247 GCTCCGCGAGCGGTAGCACCGAGTCGGGCGAGCAGCAGCCCCCGGCGCGGAGAGACAA 306  
Qy 121 AGCATGGAGGACACAAGAAATGGGAGGAAGCGGACTCTCGGGAACCTTCATTCTTCAGT 180  
Db 307 AGCATGGAGGACACAAGAAATGGGAGGAAGCGGACTCTTCAGGAACCTTCATTCTTCAGT 366  
Qy 181 GGTTTATGTTGATTTGCATTGCTGGCGGTCTGGACATCTGTAGCTGTCTGTTGGTTTGTATC 240  
Db 367 GGTTTATGTTGATTTGCATTGCTGGCGGTCTGGACATCTGTAGCTGTCTGTTGGTTTGTATC 426  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGAAAATCTAGAAATCTATGATGCTGATGCTGATGAG 300  
Db 427 TTGTTGACTATGAGGAAGTTCTAGAAAATCTAGAAATCTATGATGCTGATGCTGATGAG 486  
Qy 301 ATTTTGTGTTGATGATGCCAAAGTTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
Db 487 ATTTTGTGTTGATGATGCCAAAGTTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC 546  
Qy 361 CAGCAGTCCCCCGCAGAGAGGCTGAGCCACACACTGAGCCCCGAGAGCAGGTTCTCTGTGG 420  
Db 547 CAGCAGTCCCCCGCAGAGAGGCTGAGCCACACACTGAGCCCCGAGAGCAGGTTCTCTGTGG 606  
Qy 421 AGGCAGAACCCCGAGAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 607 AGGCAGAACCCCGAGAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 666  
Qy 481 AAATGGTACACGCAGACATG 501





\* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

\* 1 130383: contig of 130383 bp in length  
 \* 130384 130483: gap of 100 bp  
 \* 130484 132868: contig of 2385 bp in length.

#### FEATURES

source  
 1. .132868  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="RP11-23M4"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1. .130383  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left  
 130484. .132868  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right"

#### misc\_feature

#### misc\_feature

#### ORIGIN

Query Match 8.2%; Score 190; DB 2; Length 132868;  
 Best Local Similarity 100.0%; Pred. No. 2,7e-92;  
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2135 CAGGACCTGGGAGGACGAGGTCATCTTTGATGACCTCTTTGAGCAGGATG 2194  
 Db 19703 CAGGACCTGGGAGGAGCAGGAGGTCATCTTTGATGACCTCTTTGAGCAGGATG 19762  
 Qy 2195 GCAGGATCCCTCATCTTTCCGGCTGATATTCATCGTGGATGTGGCATCCGGAATGAC 2254  
 Db 19763 GCAGGATCCCTCATCTTTCCGGCTGATATTCATCGTGGATGTGGCATCCGGAATGAC 19822  
 Qy 2255 ACCACAGCAGAGCAGCGACCTTCAGCAATTTAGCATGAATTCATGCAAGCTGGGAAC 2314  
 Db 19823 ACCACAGCAGAGCAGCGACCTTCAGCAATTTAGCATGAATTCATGCAAGCTGGGAAC 19882  
 Qy 2315 TCTGGAGAGA 2324  
 Db 19883 TCTGGAGAGA 19892

#### RESULT 20

AC090094  
 LOCUS 158404 bp DNA linear PRI 15-MAR-2002  
 DEFINITION Homo sapiens chromosome 8, clone RP11-252C19, complete sequence.  
 AC090094  
 VERSION AC090094.5 GI:19482342  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 158404)  
 Birren,B., Linton,L., Nussbaum,C. and Lander,E.

Homo sapiens chromosome 8, clone RP11-252C19.  
 Unpublished

2 (bases 1 to 158404)

Birren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,S.,  
 Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,  
 Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,  
 Collumore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,  
 Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,  
 Lehoczy,J., Levine,R., Liu,G., Maclean,C., MacDonald,P.,

Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,  
 McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,  
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
 Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,  
 Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,  
 Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,  
 Souney,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Severy,P.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Velei,R., Vo,A.,  
 Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
 Zembek,L., Zimmer,A. and Zody,M.

#### Direct Submission

Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 158404)

#### JOURNAL

#### REFERENCE

#### AUTHORS

Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collumore,A.,  
 Cooke,P., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,  
 Landers,T., Lehoczy,J., Levine,R., Major,J., Marquis,N., Meneus,L.,  
 Maclean,C., MacDonald,P., McKernan,K., Meldrim,J., Mlenga,V.,  
 McCarty,M., McEwan,P., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
 Mihova,T., Mlenga,V., Murphy,T., O'Connor,T., O'Donnell,P., O'Neil,D.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Souney,C., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Velei,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

#### Direct Submission

Submitted (15-MAR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 15, 2002 this sequence version replaced gi:14626341.

#### COMMENT

All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L12758

Center clone name: 252\_C\_19

#### FEATURES

#### source

1. .158404  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8"  
 /clone="RP11-252C19"  
 /clone\_lib="RPC1-11 Human Male BAC"

733. .864

/rpt\_family="L2"

complement(1037..1417)

/rpt\_family="MLT2D"

complement(1697..1825)

/rpt\_family="LIME"

1826..2120

/rpt\_family="AluSx"

1960..1969

/note="<30 qual SNGL region"

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### unsure





```
repeat_region /rpt_family="LIPA7"
12747..13010
repeat_region /rpt_family="L2"
13177..13481
repeat_region /rpt_family="L2"
13534..13975
repeat_region /rpt_family="MER74B"
14732..14791
repeat_region /rpt_family="AT_rich"
14932..15200
repeat_region /rpt_family="LIPA5"
15260..15283
repeat_region /rpt_family="(A)n"
complement(15380..15753)
repeat_region /rpt_family="THE1C"
complement(15754..16008)
repeat_region /rpt_family="THE1C-int"
17770..17789
repeat_region /rpt_family="(A)n"
complement(17790..18318)
repeat_region /rpt_family="THE1C-int"
complement(18330..18800)
repeat_region /rpt_family="MLT1D"
19720..19754
repeat_region /rpt_family="AT_rich"
21111..21136
repeat_region /rpt_family="(TTA)n"
21150..21420
repeat_region /rpt_family="AluSx"
complement(21498..21590)
repeat_region /rpt_family="MIR"
complement(22454..22668)
repeat_region /rpt_family="FLAM_C"
23251..23447
repeat_region /rpt_family="L2"
23670..23859
repeat_region /rpt_family="MIR"
23903..24114
repeat_region /rpt_family="L2"
25384..25405
repeat_region /rpt_family="AT_rich"
25765..25893
repeat_region /rpt_family="MIR3"
26435..26587
repeat_region /rpt_family="L2"
26601..27119
repeat_region /rpt_family="LIME1"
28411..28450
repeat_region /rpt_family="(TG)n"
28613..28761
repeat_region /rpt_family="L2"
complement(29567..29672)
repeat_region /rpt_family="MIR"
29925..30025
repeat_region /rpt_family="MIR"
32387..32722
repeat_region /rpt_family="MER44D"
33432..33503
repeat_region /rpt_family="LIMEC"
33517..33596
repeat_region /rpt_family="HAL1"
33606..33694
repeat_region /rpt_family="LIM2"
35095..35226
repeat_region /rpt_family="FLAM_C"
35227..37074
repeat_region /rpt_family="LIM2"
37075..37345
repeat_region /rpt_family="LIPA5"
37346..38535
repeat_region /rpt_family="LIM2"
38522..39159
repeat_region /rpt_family="LIMEC"
39160..39429
/rpt_family="AluSg"
39435..40461
/rpt_family="L1MEC"
40464..40705
/rpt_family="L1ME1"
40771..41078
/rpt_family="AluSx"
complement(41489..41585)
repeat_region /rpt_family="L2"
complement(41844..42201)
repeat_region /rpt_family="L2"
42278..42299
/rpt_family="AT_rich"
complement(43251..43390)
repeat_region /rpt_family="MIR"
44831..44883
/rpt_family="L2"
46610..46683
/rpt_family="MERSB"

Query Match 6.7%; Score 155; DB 9; Length 186676;
Best Local Similarity 100.0%; Pred.No. 5.4e-73;
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1159 AGTGTGAGGATGATTTGGCTGAGAGAGGAGGAATATAGGTTGCTAGTGGAGCCATCG 1218
|||||
Db 5180 AGTGTGAGGATGATTTGGCTGAGAGAGGAGGAATATAGGTTGCTAGTGGAGCCATCG 5121

QY 1219 AGACCTACCAAGAGTGGCCAGCTTACTGATGTCCTGACAGCTGCTGAAGCTGAGTT 1278
|||||
Db 5120 AGACCTACCAAGAGTGGCCAGCTTACTGATGTCCTGACAGCTGCTGAAGCTGAGTT 5061

QY 1279 TGAAGCGTGTCTCAGACAGGCAACAATTTCTAGGT 1313
|||||
Db 5060 TGAAGCGTGTCTCAGACAGGCAACAATTTCTAGGT 5026

RESULT 23
CO657781 191 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 2707 from Patent WO02070737.
DEFINITION CO657781
ACCESSION CO657781
VERSION CO657781.1 GI:42121483
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.B. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 2707 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source
1..191
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 5.8%; Score 135; DB 6; Length 191;
Best Local Similarity 100.0%; Pred.No. 5.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 TCCGCCAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGGTTCTCTGTGGAGGAG 426
|||||
Db 1 TCCGCCAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGGTTCTCTGTGGAGGAG 60

QY 427 AACCCAGAAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCCATGAATGG 486
|||||
Db 61 AACCCAGAAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCCATGAATGG 120
```













```

ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 1777)
AUTHORS
Jones, L.R., Zhang, L., Sanborn, K., Jorgensen, A.O. and Kelley, J.
TITLE
Purification, primary structure, and immunological characterization
of the 26-kDa calsequestrin binding protein (junctin) from cardiac
junctional sarcoplasmic reticulum
J. Biol. Chem. 270 (51), 30787-30796 (1995)
JOURNAL
MEDLINE
96107245
PUBMED
8530521
REFERENCE
2 (bases 1 to 1777)
AUTHORS
Jones, L.R., Zhang, L., Sanborn, K., Jorgensen, A.O. and Kelley, J.
TITLE
Direct Submission
Submitted (12-OCT-1995) Larry R. Jones, Medicine, Krannert Inst.
Cardiology, 1111 W. 10 Street, Indianapolis, IN 46202, USA
JOURNAL
FEATURES
source
1. .1777
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/tissue_type="heart"
89. .721
/function="calsequestrin binding protein"
/note="junctin"
/codon_start=1
/product="junctional sarcoplasmic reticulum protein"
/protein_id="AAC48501.1"
/db_xref="GI:1163913"
/translation="MAEETKGGHKNRKGSLGSSSFTWFMVALLGVWTSVAVVWF
DLVDYEVILGKLVADGGDFDVDDAKVLLEGFGGVAKRTAKVKEIKELKKE
KEKTSRKNKNEERKKGKEDKDKTADADISRKESPKGKDKREXENGLDKS
AKAKESRKSSTNVKDASSKTASDKDDTKEGKTSKTHSAKGNQKRN"
polyA_site
1777
/note="44 A nucleotides"
ORIGIN
Query Match 2.3%; Score 53; DB 4; Length 1777;
Best Local Similarity 100.0%; Pred. No. 8.6e-17;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 277 TCTATGATGCTGATGGATGGAGATTTTGATGCGATGCCCAAGTTT 329
|||||
Db 261 TCTATGATGCTGATGGATGGAGATTTTGATGCGATGCCCAAGTTT 313
|||||
RESULT 34
AF198966
LOCUS
AF198966
DEFINITION
Oryctolagus cuniculus cardiac junctin mRNA, partial cds.
ACCESSION
AF198966
VERSION
AF198966.1 GI:7141074
KEYWORDS
Oryctolagus cuniculus (rabbit)
ORGANISM
Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE
1 (bases 1 to 556)
AUTHORS
Wetzel, G.T., Ding, S. and Chen, F.
TITLE
Molecular cloning of junctin from human and developing rabbit heart
Mol. Genet. Metab. 69 (3), 252-258 (2000)
JOURNAL
MEDLINE
20232172
PUBMED
10767180
REFERENCE
2 (bases 1 to 556)
AUTHORS
Wetzel, G.T., Ding, S. and Chen, F.
TITLE
Direct Submission
Submitted (26-OCT-1999) Pediatrics Cardiology, UCLA School of
Medicine, 675 Charles E. Young Drive South, 3754 MEL Building, Los
Angeles, CA 90095-7045, USA
JOURNAL
FEATURES
source
1. .556
/organism="Oryctolagus cuniculus"
/mol_type="mRNA"

```

```

GDCFDVDDAKVLLGLKERSPSERTPPAEATHAELEAOPEGADIONVEDEVKEQIO
SLIQSVHTDHDLEADGLAGEPOVEDEFLTVTDDDRFEDLEPGTVHERIEDTYHVE
DTASQNHNDMEEMTNEQNSDPSEAVTDAGVLLPHAEVRHQDVPVPESEHGV
RKSIDNTIDDSSTIGIEINVASVEEQDTPPVKKKPKLLNKFDXTIKAEIDAAELRK
RGIKEAVNAFEELVRKYPOSPRARYGKACQEDDLAERQSNVRLAETIYEAADL
PAPDLDVLLKRSRQOFLGHMRGSLTLQRLVOLFPDSTTLKNDLGVGLLGLD
NDSKVVYEVSLVNTDNDGFAKHGFIILAKQNKISEIPFIKESIGESGPDGDDGF
VHLGDAMORVGNKEAYKVELGHKGHFASVWQSLNVNGLKAQPMWTPRETGYE
LVKSLERWKLIRDEGLVMDKAGLFLPEDENLREKGDWSQFTLWQOGRKNNAECK
AKTICALLEKSETITGCRGQIKYSIMHFGTHVPHGTNTCLRMLHGLVIFPECKG
ICANETITWEEGKVLIFDDSFHEVWQDASSFRLLIFVDVWHPELTPQRRSLPAI"
2395..4419
/gene="Asph"
3'UTR
polyA_signal
ORIGIN
Query Match 1.5%; Score 35; DB 10; Length 4419;
Best Local Similarity 100.0%; Pred. No. 7.3e-07; Indels 0; Gaps 0;
Matches 35; Conservative 0; Mismatches 0;
QY 2208 TCCTTCGGCTGATATTCATCGTGGATGTGGCA 2242
|||||
DB 2314 TCCTTCGGCTGATATTCATCGTGGATGTGGCA 2348
|||||

RESULT 36
H010058511 6024 bp DNA linear ROD 13-MAR-2001
LOCUS Mus musculus aspartyl beta-hydroxylase (ASPH) gene, exon 24 and
DEFINITION complete cds.
ACCESSION AF289215
VERSION AF289215.2 GI:13310928
KEYWORDS
SEGMENT
SOURCE 11 of 11
MUS musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Link, J., O'Neill, K.T., Focht, R.J., Scully, M.S., Hollis, J.M.,
Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,
Hollis, G.F. and Friedman, P.A.
Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved
isoform of Asph missing the catalytic domain share exons with
junctin
J. Biol. Chem. 275 (50), 39543-39554 (2000)
MEDLINE 20564328
PUBMED 10956665
REFERENCE 2 (bases 1 to 6024)
Henderson, N.L., Dinchuk, J.E., Burn, T.C., Hollis, G.F. and
Friedman, P.A.
Direct Submission
Submitted (25-JUL-2000) Applied Biotechnology, DuPont
Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA
3 (bases 1 to 6024)
Henderson, N.L., Dinchuk, J.E., Burn, T.C., Hollis, G.F. and
Friedman, P.A.
Direct Submission
Submitted (13-MAR-2001) Applied Biotechnology, DuPont
Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA
REMARK Sequence update by submitter
COMMENT On Mar 13, 2001 this sequence version replaced gi:11692644.
FEATURES
source
1..6024
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
order(AF289205.1:1772..2483,AF289206.1:1..1951,
AF289207.2:1..31245,AF289208.1:1..2633,AF289209.1:1..2013,
AF289210.1:1..5161,AF289211.1:1..6629,AF289212.1:1..2335,
AF289213.1:1..5657,AF289214.1:1..12095,1..>6000)
/gene="ASPH"

join(AF289205.1:1772..2075,AF289206.1:1105..1173,
AF289207.2:1425..1514,AF289207.2:4438..4512,
AF289207.2:7694..7807,AF289207.2:9820..9852,
AF289207.2:10948..11004,AF289207.2:12520..12567,
AF289207.2:12997..13029,AF289207.2:15883..15921,
AF289207.2:16600..16635,AF289207.2:20179..20223,
AF289208.1:1520..1605,AF289209.1:292..378,
AF289210.1:2354..2504,AF289210.1:4279..4415,
AF289211.1:1764..1862,AF289211.1:4834..4923,
AF289212.1:428..565,AF289213.1:2049..2184,
AF289214.1:5360..5451,AF289214.1:5805..5938,1665..>6000)
/gene="ASPH"
/product="aspartyl beta-hydroxylase"
join(AF289205.1:1946..2075,AF289206.1:1105..1173,
AF289207.2:1425..1514,AF289207.2:4438..4512,
AF289207.2:7694..7807,AF289207.2:9820..9852,
AF289207.2:10948..11004,AF289207.2:12520..12567,
AF289207.2:12997..13029,AF289207.2:15883..15921,
AF289207.2:16600..16635,AF289207.2:20179..20223,
AF289208.1:1520..1605,AF289209.1:292..378,
AF289210.1:2354..2504,AF289210.1:4279..4415,
AF289211.1:1764..1862,AF289211.1:4834..4923,
AF289212.1:428..565,AF289213.1:2049..2184,
AF289214.1:5360..5451,AF289214.1:5805..5938,1665..1815)
/gene="ASPH"
/codon_start=1
/product="aspartyl beta-hydroxylase"
/protein_id="AG339613.1"
/db_xref="GI:11692645"
/translation="MAPRKNAGKGGSSSSGSGSGSGSPSTGSGSSSSSPFGARRG
KLVYDADGQDFDVAADAKVLLGLKERSPSERTPPAEATHAELEAOPEGADIONVE
DEVKEQIQSLLESVHTDHDLEADGLAGEPOVEDEFLTVTDDDRFEDLEPGTVHER
IEDTYHVEDTASQNHNDMEEMTNEQNSDPSEAVTDAGVLLPHAEVRHQDVPVPE
EPSEHGVESIDNTIDDSSTIGIEINVASVEEQDTPPVKKKPKLLNKFDXTIKAEI
DAAKLRKGIKEAVNAFEELVRKYPOSPRARYGKACQEDDLAERQSNVRLAETIYEA
TYQEAADLPDAPDLDVLLKRSRQOFLGHMRGSLTLQRLVOLFPDSTTLKNDLGLD
VGVLGDNDKSAKVVYEVSLVNTDNDGFAKHGFIILAKQNKISEIPFIKESIGESG
PGTDDGRFVHLGDAMORVGNKEAYKVELGHKGHFASVWQSLNVNGLKAQPMWTPRE
TGYEYLVKSLERWKLIRDEGLVMDKAGLFLPEDENLREKGDWSQFTLWQOGRKNNA
CKGKVNACKGAPKTCALLEKSETITGCRGQIKYSIMHFGTHVPHGTNTCLRMLHGL
VIPECKGKIRCANETRTWEEGKVLIFDDSFHEVWQDASSFRLLIFVDVWHPELTPQ
RRSLPAI"
1665..1815
/gene="ASPH"
number=24
3801..3806
/gene="ASPH"
5995..6000
/gene="ASPH"

exon
polyA_signal
polyA_signal
ORIGIN
Query Match 1.5%; Score 35; DB 10; Length 6024;
Best Local Similarity 100.0%; Pred. No. 7.3e-07; Indels 0; Gaps 0;
Matches 35; Conservative 0; Mismatches 0;
QY 2208 TCCTTCGGCTGATATTCATCGTGGATGTGGCA 2242
|||||
DB 1735 TCCTTCGGCTGATATTCATCGTGGATGTGGCA 1769
|||||

RESULT 37
AF289487
LOCUS Mus musculus aspartyl beta-hydroxylase 6.6 kb transcript (Asph)
DEFINITION mRNA, complete cds; alternatively spliced.
ACCESSION AF289487
VERSION AF289487.1 GI:11878111
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 6629)
```

**AUTHORS** Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P., Link, J., O'Neill, K.T., Focht, R.J., Scully, M.S., Hollis, J.M., Hollis, G.F. and Friedman, P.A.  
**TITLE** Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junctin  
**JOURNAL** J. Biol. Chem. 275 (50), 39543-39554 (2000)  
**MEDLINE** 20564328  
**PUBMED** 10956665  
**REFERENCE** 2 (bases 1 to 6629)  
**AUTHORS** Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and Friedman, P.A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA  
**FEATURES**  
 source  
 1..6629  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /chromosome="8"  
 /map="8q12 at STS WT-11767"  
 /tissue\_type="liver"  
 1..6629  
 /gene="Asph"  
 1..174  
 /gene="Asph"  
 175..2394  
 /gene="Asph"  
 /note="BAH; asparaginyl beta-hydroxylase; Asph; alternatively spliced"  
 /codon\_start=1  
 /product="aspartyl beta-hydroxylase 6.6 kb transcript"  
 /protein\_id="AAG40809.1"  
 /db\_xref="GI:11878112"  
 /translation="MAPRKNAKGGGSSSGSGSGSGSPSTGSSSSSPGARRE AKHGDKVDDAKVLLGKERSPSRTFPPEAEATHAELEQAPGADIQNVDEKVEIQI QGDFDNDADKGLKERSPSRTFPPEAEATHAELEQAPGADIQNVDEKVEIQI SLLQSVTHDLEADGLAGEPOVEDFLVTDSDDREDFLEPTVHEETDTHVVE DTASQHPNDMEETNEQNSDPSAVTDAGVLLPHAEVVRHODYDEPVRSHEGV EISDNTDSSIISEINVASVEEQDTPPVKKKPKLLNKFDKTIKAELEDAEKLKX RGIKEAVNAFELVKRYQSPRARKAQCEDDIAEKQNEVLRLRAETIYQAAALD PDAPTDLVKLSLRKRSERQFLGHRGSLTLRLIVQLFPSPDITLKNLGVGYLLGD NDSAKVYEVNLNVTNDGFAKVHYGFLKAKNKISSEIPIYKIEGIESGDPGTDGRF YHLCGMQVGNKAYKWEYELGHRGHFASVMQSLVNVGLKAPQWMTPRETCYTE LVKSLRNKLLRDGLVMDKAGLFLPEDENLEKGDWSQFTLWQQGKRNENACKG APKCALLLEKSETTGCRGQIKYSIMHPGTHVWHPGTNCRLEMHGLGLVLPKEGCK IRCANETREEGKVLIFDDSPFEHEVWQDASSFRULIFIVDVVHPELTPOQRSLPAL"  
 2395..6629  
 /gene="Asph"  
 4379..4384  
 /gene="Asph"  
 6573..6578  
 /gene="Asph"  
 3'UTR  
 polyA\_signal  
 polyA\_signal  
**ORIGIN**  
 Query Match 1.5%; Score 35; DB 10; Length 6629;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2208 TCCTTCGGGTGATATTCATCGTGGATGTGGCA 2242  
 |||||  
 DB 2314 TCCTTCGGGTGATATTCATCGTGGATGTGGCA 2348  
 |||||  
**RESULT 38**  
**LOCUS** AL671970/c 168569 bp DNA linear ROD 29-APR-2004  
**DEFINITION** Mouse DNA sequence from clone RP23-152A20 on chromosome 4, complete sequence.  
**ACCESSION** AL671970  
**VERSION** AL671970.8 GI:46878903

**KEYWORDS** HTG.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus  
**REFERENCE** Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 168569)  
**AUTHORS** Pelan, S.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (29-APR-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
**COMMENT** humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Apr 29, 2004 this sequence version replaced gi:21531211.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquerry@sanger.ac.uk  
 -----  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
 RP23-152A20 is from the RPI-23 Mouse BAC Library constructed by the group of Pieter de Jong.  
 For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBACe3.6  
 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.  
**FEATURES**  
 source  
 1..168569  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-152A20"  
 /clone\_lib="RPCI-23"  
**ORIGIN**  
 Query Match 1.5%; Score 35; DB 10; Length 168569;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2208 TCCTTCGGGTGATATTCATCGTGGATGTGGCA 2242  
 |||||  
 DB 73512 TCCTTCGGGTGATATTCATCGTGGATGTGGCA 73478  
 |||||  
**RESULT 39**  
**LOCUS** AF289491 524 bp mRNA linear ROD 18-DEC-2000  
**DEFINITION** Mus musculus aspartyl beta-hydroxylase cardiac isoform 1 (Asph)  
**ACCESSION** AF289491  
**VERSION** AF289491.1 GI:11878119  
**KEYWORDS** Mus musculus (house mouse)  
**SOURCE**

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 524)  
AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P., Link, J., O'Neil, K.T., Pocht, R.J., Scully, M.S., Hollis, J.M., Hollis, G.F. and Friedman, P.A.  
TITLE Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junction  
JOURNAL J. Biol. Chem. 275 (50), 39543-39554 (2000)  
MEDLINE 20564328  
PUBMED 10956665  
REFERENCE 2 (bases 1 to 524)  
AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and Friedman, P.A.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA  
FEATURES  
source  
1..524  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/map="8q12 at STS WT-11767"  
/tissue type="heart"  
1..>524  
/gene="Asph"  
1..124  
/gene="Asph"  
125..>524  
/gene="Asph"  
/note="asparaginyl beta-hydroxylase; Asph; alternatively spliced"  
/codon\_start=1  
/product="aspartyl beta-hydroxylase cardiac isoform 1"  
/protein\_id="AAG40813.1"  
/db\_xref="GI:11878120"  
/translation="MAEDKEAKHGKNGRRGGISGGSFFTFWVIALGLVWTSVAVV  
WFDLVDVEYVLGKGYDADGDGDFDDAKVLLGLKERSPSTRTPEAETHALEE  
QAPEGADIONVEDEVKEIQSLQSSVHTDH"  
ORIGIN  
Query Match 1.4%; Score 32; DB 10; Length 524;  
Best Local Similarity 100.0%; Pred. No. 3.2e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 303 TTGTGATGGATGATGCCAAAGTTTATTAGG 334  
|||||  
Db 329 TTGTGATGGATGATGCCAAAGTTTATTAGG 360  
|||||  
RESULT 40  
BC061098  
LOCUS BC061098 1763 bp mRNA linear ROD 04-NOV-2003  
DEFINITION Mus musculus cDNA clone MGC:74216 IMAGE:6533192, complete cds.  
ACCESSION BC061098  
VERSION BC061098.1 GI:38173948  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1763)  
AUTHORS Klausner, R.D., Collins, F.S., Wagner, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L.H., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, S.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932  
2 (bases 1 to 1763)  
Strausberg, R.  
Direct Submission  
Submitted (03-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Michael Brownstein / Ted Usdin  
Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: TRAL Plate: 53 Row: a Column: 10  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 12746447.  
FEATURES  
Location/Qualifiers  
1..1763  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/Clone="MGC:74216 IMAGE:6533192"  
/tissue\_type="Brain, mouse"  
/clone\_lib="NIH MGC\_144"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"  
199..837  
/codon\_start=1  
/product="Unknown (protein for MGC:74216)"  
/protein\_id="AAH61098.1"  
/db\_xref="GI:38173948"  
/translation="WAPKNAKGGGSSSGSGSGSGSPSTGSSSGSPGARRE  
AKGCHNGRGGISGGSFFTFWVIALGLVWTSVAVVFDLVDHVLAKADFRYN  
LSEVLOGKLVGDADGDGDFDDAKVLLGLKERSIDSEEVLTILAESDWF  
YGFSLFDIMTFPEMEEESEETADGVDGTSQNEGVQGVTCVILDHNO"  
385..618  
/note="Asp-B-Hydro\_N; Region: Aspartyl beta-hydroxylase N-terminal region. This family includes the N-terminal regions of the junction, junctate and aspartyl beta-hydroxylase proteins. Junctate is an integral ER/SR membrane calcium binding protein, which comes from an alternatively spliced form of the same gene that generates aspartyl beta-hydroxylase and junctin. Aspartyl beta-hydroxylase catalyses the post-translational hydroxylation of aspartic acid or asparagine residues contained within epidermal growth factor (EGF) domains of proteins"



```

ORIGIN
  /db_xref="CDD:pfam05279"

Query Match      1.4%; Score 32; DB 10; Length 1763;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAGG 334
Db 562 TTTGATGTGGATGATGCCAAAGTTTATTAGG 593

RESULT 41
H010058S02      1951 bp      DNA      linear      ROD 13-MAR-2001
LOCUS
DEFINITION      Mus musculus aspartyl beta-hydroxylase (ASPH) gene, exon 3.
ACCESSION      AF289206
VERSION      AF289206.1 GI:11692635
KEYWORDS
SEGMENT
SOURCE      Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,
Link, J., O'Neill, K.I., Focht, R.J., Scully, M.S., Hollis, J.M.,
Hollis, G.F. and Friedman, P.A.
Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved
isoform of Asph missing the catalytic domain share exons with
junctin
J. Biol. Chem. 275 (50), 39543-39554 (2000)
MEDLINE
PUBMED
REFERENCE
AUTHORS
Henderson, N.L., Dinchuk, J.E., Burn, T.C., Hollis, G.F. and
Friedman, P.A.
Direct Submission
Submitted (25-JUL-2000) Applied Biotechnology, DuPont
Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA
FEATURES
source
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  1105..1173
  /gene="ASPH"
  /number=3
exon

ORIGIN
Query Match      1.4%; Score 32; DB 10; Length 1951;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAGG 334
Db 1143 TTTGATGTGGATGATGCCAAAGTTTATTAGG 1174

RESULT 42
AF302653
LOCUS
DEFINITION      Mus musculus cardiac junctate 1 mRNA, complete cds.
ACCESSION      AF302653
VERSION      AF302653.2 GI:16507230
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
Hong, C.-S. and Kim, D.H.
Cloning of mouse junctin homologs
TITLE

JOURNAL
REFERENCE
AUTHORS
Hong, C.-S. and Kim, D.H.
Direct Submission
Submitted (05-SEP-2000) Life Science, Kwangju Institute of Science
and Technology, Puk-gu, Oryong-dong, Kwangju 500-712, South Korea
COMMENT
On Oct 29, 2001 this sequence version replaced gi:15824389.
FEATURES
source
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="BALB/c"
  /db_xref="taxon:10090"
  67..879
  /codon_start=1
  /product="cardiac junctate 1"
  /protein_id="AA09319.2"
  /db_xref="GI:16507231"
  /translation="MAEDKEAKHGHNRRGSGSFTFMVTLGLGVWTSVAVV
WFDLVYEEVLGKLVYDADGDGDFVDKALLGLKRSPPSRTFFPFAETHAELEE
QAPGADIQNVDEVEKEIQSLIQESVHTDHLADGLAGEPQVEFDLTVTSDDR
FEDLEPTGTHEETEDTVHVEDTASQNHNDMEEMTNEQNSDSEAVTAGVLLPHAE
EVRHQDYDEPVYEPSEHGVEISDNTIDDSIISEINVASVEEQQDTDPDT"
ORIGIN
Query Match      1.4%; Score 32; DB 10; Length 2617;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAGG 334
Db 271 TTTGATGTGGATGATGCCAAAGTTTATTAGG 302

RESULT 43
BC015281
LOCUS
DEFINITION      Mus musculus aspartate-beta-hydroxylase, mRNA (cDNA clone
IMAGE:4235968), complete cds.
ACCESSION      BC015281
VERSION      BC015281.1 GI:15929715
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
MEDLINE
PUBMED
REFERENCE
AUTHORS
Strausberg, R.
Direct Submission
Submitted (bases 1 to 2770)
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Strausberg, R.
Direct Submission
Submitted (bases 1 to 2770)
TITLE

```



REFERENCE 1 (bases 1 to 198991)  
AUTHORS Lovell, J.  
TITLE Direct Submission  
JOURNAL Submitted (01-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hamquerry@sanger.ac.uk  
COMMENT On Oct 2, 2002 this sequence version replaced gi:23395448.  
----- Genom Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: hamquerry@sanger.ac.uk  
-----  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at  
http://www.sanger.ac.uk/projects/C\_elegans/wormpep RP23-70D2 is from the RPI-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see http://www.chori.org/bacpac/home.htm  
VECTOR: pBACE3.6.  
Location/Qualifiers  
1..198991  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone\_lib="RP23-70D2"  
/clone\_lib="RPI-23"

source  
1..198991  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone\_lib="RP23-70D2"  
/clone\_lib="RPI-23"

ORIGIN  
Query Match 1.4%; Score 32; DB 10; Length 198991;  
Best Local Similarity 100.0%; Pred. No. 3.4e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 303 TTGTGATGCGATGATGCCAAAGTTTATTAGG 334  
DB 20973 TTGTGATGCGATGATGCCAAAGTTTATTAGG 20942  
|||||  
RESULT 46  
AF289492 566 bp mRNA linear ROD 18-DEC-2000  
LOCUS Mus musculus asparaginyl beta-hydroxylase cardiac isoform 2 mRNA, partial cds; alternatively spliced.  
DEFINITION AF289492  
ACCESSION AF289492  
VERSION AF289492.1 GI:11878121  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 566)  
REFERENCE Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P., Link, J., O'Neil, K.T., Focht, R.J., Scully, M.S., Hollis, J.M., Hollis, G.F., and Friedman, P.A.  
AUTHORS Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with

junctin J. Biol. Chem. 275 (50), 39543-39554 (2000)  
MEDLINE 205643328  
PUBMED 10956665  
REFERENCE 2 (bases 1 to 566)  
AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and Friedman, P.A.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA  
FEATURES  
source  
1..566  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/chromosome="8"  
/map="8q12 at STS WT-11767"  
/tissue\_type="heart"  
1..124  
125..566  
/note="BAH; asparaginyl beta-hydroxylase; Asph; alternatively spliced"  
/codon\_start=1  
/product="aspartyl beta-hydroxylase cardiac isoform 2"  
/protein\_id="AAC40814.1"  
/db\_xref="GI:11878122"  
/translation="MAEDKEAKHGKNGRGISGGSFTTFWMTALGLGWTSAVW WFLDVEYVLGVLGVADGDGDFVDDAKVLGFGGLAKFKTKAKGLKERSPSER TFPPEAAHAELEBEQAPEGADIQNVEDEVEKEIQSLLESVHTDH"  
ORIGIN  
QY 303 TTGTGATGCGATGATGCCAAAGTTTATTAG 333  
DB 329 TTGTGATGCGATGATGCCAAAGTTTATTAG 359  
|||||  
RESULT 47  
AF221854 919 bp mRNA linear ROD 02-FEB-2001  
LOCUS Mus musculus cardiac junctin mRNA, complete cds.  
DEFINITION AF221854  
ACCESSION AF221854  
VERSION AF221854.1 GI:12655824  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 919)  
REFERENCE Jones, L.R., Zhang, L., Sanborn, K., Jorgensen, A.O. and Kelley, J.  
AUTHORS Purification, primary structure, and immunological characterization of the 26-kDa calsequestrin binding protein (junctin) from cardiac junctional sarcoplasmic reticulum  
TITLE J. Biol. Chem. 270 (51), 30787-30796 (1995)  
JOURNAL 96107245  
MEDLINE 8530521  
PUBMED 8530521  
REFERENCE 2 (bases 1 to 919)  
AUTHORS Zhang, L. and Jones, L.R.  
TITLE Direct Submission  
JOURNAL Submitted (06-JAN-2000) 47, Sigma Chemical Co., N-3, 3300 South Second Street, St. Louis, MO 63118, USA  
FEATURES  
source  
1..919  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="cardiac muscle"  
143..766  
/note="calsequestrin binding protein"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 347 TTTGATGTGGATGATGCCAAAGTTTATTAG 377

RESULT 48  
 AF223414  
 LOCUS Mus musculus junctin-2 mRNA, complete cds. ROD 19-DEC-2002  
 DEFINITION Mus musculus junctin-2 mRNA, complete cds.  
 ACCESSION AF223414  
 VERSION AF223414.1 GI:27261147  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Hong, C.-S. and Kim, D.H.  
 TITLE Mouse junctin-2 mRNA  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 988)  
 AUTHORS Hong, C. and Kim, D.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JAN-2000) Life Science, Kwangju Institute of Science and Technology (K-JIST), Puk-gu, Oryong-dong, Kwangju 500-712, South Korea

FEATURES  
 source  
 1..988  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALE/c"  
 /db\_xref="taxon:10090"  
 /tissue\_type="heart"  
 64..801  
 /codon\_start=1  
 /product="junctin-2"  
 /protein\_id="AA087550.1"  
 /db\_xref="GI:27261148"  
 /translation="MAPRNNAKGSGNSGSGSGSGSPSTGSSGSSSPGARRE  
 AKGGHNNRGGISGGSGFTFWVIALGVWTSVAVWFDLYEVLGKLVYDAD  
 GDGDFDVAKVLGGPGGLAKRTKAKKEPIKEELKRGKAVPSKNEERQCKE  
 QEDRGKRPDSDTSQKASAAAGKDRDRDKKSSKSKSVKAVETKAVSSKVA  
 ARDKDRGRSSSGHARVSKENGQKRN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 347 TTTGATGTGGATGATGCCAAAGTTTATTAG 377

RESULT 48  
 AF223414  
 LOCUS Mus musculus junctin-2 mRNA, complete cds. ROD 19-DEC-2002  
 DEFINITION Mus musculus junctin-2 mRNA, complete cds.  
 ACCESSION AF223414  
 VERSION AF223414.1 GI:27261147  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Hong, C.-S. and Kim, D.H.  
 TITLE Mouse junctin-2 mRNA  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 988)  
 AUTHORS Hong, C. and Kim, D.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JAN-2000) Life Science, Kwangju Institute of Science and Technology (K-JIST), Puk-gu, Oryong-dong, Kwangju 500-712, South Korea

FEATURES  
 source  
 1..988  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALE/c"  
 /db\_xref="taxon:10090"  
 /tissue\_type="heart"  
 64..801  
 /codon\_start=1  
 /product="junctin-2"  
 /protein\_id="AA087550.1"  
 /db\_xref="GI:27261148"  
 /translation="MAPRNNAKGSGNSGSGSGSGSPSTGSSGSSSPGARRE  
 AKGGHNNRGGISGGSGFTFWVIALGVWTSVAVWFDLYEVLGKLVYDAD  
 GDGDFDVAKVLGGPGGLAKRTKAKKEPIKEELKRGKAVPSKNEERQCKE  
 QEDRGKRPDSDTSQKASAAAGKDRDRDKKSSKSKSVKAVETKAVSSKVA  
 ARDKDRGRSSSGHARVSKENGQKRN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 988;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 382 TTTGATGTGGATGATGCCAAAGTTTATTAG 412

RESULT 49  
 AF223413  
 LOCUS Mus musculus junctin-1 mRNA, complete cds. ROD 19-DEC-2002  
 DEFINITION Mus musculus junctin-1 mRNA, complete cds.  
 ACCESSION AF223413

FEATURES  
 source  
 1..988  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALE/c"  
 /db\_xref="taxon:10090"  
 /tissue\_type="heart"  
 64..801  
 /codon\_start=1  
 /product="junctin-2"  
 /protein\_id="AA087550.1"  
 /db\_xref="GI:27261148"  
 /translation="MAPRNNAKGSGNSGSGSGSGSPSTGSSGSSSPGARRE  
 AKGGHNNRGGISGGSGFTFWVIALGVWTSVAVWFDLYEVLGKLVYDAD  
 GDGDFDVAKVLGGPGGLAKRTKAKKEPIKEELKRGKAVPSKNEERQCKE  
 QEDRGKRPDSDTSQKASAAAGKDRDRDKKSSKSKSVKAVETKAVSSKVA  
 ARDKDRGRSSSGHARVSKENGQKRN"

AF223413.1 GI:27261145  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Hong, C. and Kim, D.H.  
 TITLE Mouse junctin-1 mRNA  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1013)  
 AUTHORS Hong, C. and Kim, D.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JAN-2000) Life Science, Kwangju Institute of Science and Technology (K-JIST), Puk-gu, Oryong-dong, Kwangju 500-712, South Korea

FEATURES  
 source  
 1..1013  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALE/c"  
 /db\_xref="taxon:10090"  
 /tissue\_type="heart"  
 168..791  
 /codon\_start=1  
 /product="junctin-1"  
 /protein\_id="AA087549.1"  
 /db\_xref="GI:27261146"  
 /translation="MAEDKEAKHGHNRRGGISGGSGFTFWVIALGVWTSVAVV  
 WFDLYEVLGKLVYDADGDGDFDVAKVLGGPGGLAKRTKAKKEPIKEELK  
 KERKAVPSKNEERQCKEEDRGKRPDSDTSQKASAAAGKDRDRDKKSSDKSS  
 KSKESWKKAVETKAVSSKVAARDKDRGRSSSGHARVSKENGQKRN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 372 TTTGATGTGGATGATGCCAAAGTTTATTAG 402

RESULT 50  
 AF289490  
 LOCUS Mus musculus junctin mRNA, complete cds. ROD 18-DEC-2000  
 DEFINITION Mus musculus junctin mRNA, complete cds.  
 ACCESSION AF289490  
 VERSION AF289490.1 GI:11878117  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1733)  
 Dinchuk, J. E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,  
 Link, J., O'Neil, K.T., Focht, R.J., Scully, M.S., Hollis, J.M.,  
 Hollis, G.F. and Friedman, P.A.  
 TITLE Aspartyl beta -hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junctin  
 JOURNAL J. Biol. Chem. 275 (50), 39543-39554 (2000)  
 MEDLINE 20564328  
 PUBMED 10956665  
 REFERENCE 2 (bases 1 to 1733)  
 AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and Friedman, P.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA

FEATURES  
 source  
 1..1733  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALE/c"  
 /db\_xref="taxon:10090"  
 /tissue\_type="heart"  
 168..791  
 /codon\_start=1  
 /product="junctin-1"  
 /protein\_id="AA087549.1"  
 /db\_xref="GI:27261146"  
 /translation="MAEDKEAKHGHNRRGGISGGSGFTFWVIALGVWTSVAVV  
 WFDLYEVLGKLVYDADGDGDFDVAKVLGGPGGLAKRTKAKKEPIKEELK  
 KERKAVPSKNEERQCKEEDRGKRPDSDTSQKASAAAGKDRDRDKKSSDKSS  
 KSKESWKKAVETKAVSSKVAARDKDRGRSSSGHARVSKENGQKRN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 372 TTTGATGTGGATGATGCCAAAGTTTATTAG 402

RESULT 50  
 AF289490  
 LOCUS Mus musculus junctin mRNA, complete cds. ROD 18-DEC-2000  
 DEFINITION Mus musculus junctin mRNA, complete cds.  
 ACCESSION AF289490  
 VERSION AF289490.1 GI:11878117  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1733)  
 Dinchuk, J. E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,  
 Link, J., O'Neil, K.T., Focht, R.J., Scully, M.S., Hollis, J.M.,  
 Hollis, G.F. and Friedman, P.A.  
 TITLE Aspartyl beta -hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junctin  
 JOURNAL J. Biol. Chem. 275 (50), 39543-39554 (2000)  
 MEDLINE 20564328  
 PUBMED 10956665  
 REFERENCE 2 (bases 1 to 1733)  
 AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and Friedman, P.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA

FEATURES  
 source  
 1..1733  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALE/c"  
 /db\_xref="taxon:10090"  
 /tissue\_type="heart"  
 168..791  
 /codon\_start=1  
 /product="junctin-1"  
 /protein\_id="AA087549.1"  
 /db\_xref="GI:27261146"  
 /translation="MAEDKEAKHGHNRRGGISGGSGFTFWVIALGVWTSVAVV  
 WFDLYEVLGKLVYDADGDGDFDVAKVLGGPGGLAKRTKAKKEPIKEELK  
 KERKAVPSKNEERQCKEEDRGKRPDSDTSQKASAAAGKDRDRDKKSSDKSS  
 KSKESWKKAVETKAVSSKVAARDKDRGRSSSGHARVSKENGQKRN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 372 TTTGATGTGGATGATGCCAAAGTTTATTAG 402

```
Db 251 TTTGATGTGGATGATGCCAAAGTTTTATTAG 281

RESULT 52
AF302654 2637 bp mRNA linear ROD 29-OCT-2001
LOCUS Mus musculus cardiac junctate 2 mRNA, complete cds.
DEFINITION AF302654
ACCESSION AF302654
VERSION AF302654.1 GI:15824391
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2637)
AUTHORS Hong, C.-S. and Kim, D.H.
TITLE Cloning of mouse junctin homologs
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2637)
AUTHORS Hong, C.-S. and Kim, D.H.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2000) Life Science, Kwangju Institute of Science
and Technology, Puk-gu, Oryong-dong, Kwangju 500-712, South Korea
FEATURES
source
1..2637
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
120..899
/codon_start=1
/product="cardiac junctate 2"
/protein_id="AAL09320.1"
/db_xref="GI:15824392"
/translation="MAEDKEAKHGKNGRRGIGSGSFTFMVIALLGWTSVAVV
WFDLVYEEVLKGLGYDADGDGDFDVKVLLDGLGPGGLAKRTAKAKPIKEIK
KRGKAVPSKNEERQKQEDRGKRPDSDTSOKASAAGRDRDKEKASDSS
KSKSWKAVETKAVSSKVAARDKDRGRSSSGHGVSKENGQKRN"
3'UTR 749..1733
polyA_signal 1700..1705
ORIGIN
Query Match 1.3%; Score 31; DB 10; Length 1733;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTTATTAG 333
Db 329 TTTGATGTGGATGATGCCAAAGTTTTATTAG 359

RESULT 51
AF302655 2432 bp mRNA linear ROD 29-OCT-2001
LOCUS Mus musculus cardiac junctate 3 mRNA, complete cds.
DEFINITION AF302655
ACCESSION AF302655
VERSION AF302655.1 GI:15824393
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2432)
AUTHORS Hong, C.-S. and Kim, D.H.
TITLE Cloning of mouse junctin homologs
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2432)
AUTHORS Hong, C.-S. and Kim, D.H.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2000) Life Science, Kwangju Institute of Science
and Technology, Puk-gu, Oryong-dong, Kwangju 500-712, South Korea
FEATURES
source
1..2432
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
47..694
/codon_start=1
/product="cardiac junctate 3"
/protein_id="AAL09321.1"
/db_xref="GI:15824394"
/translation="MAEDKEAKHGKNGRRGIGSGSFTFMVIALLGWTSVAVV
WFDLVYEEVLKGLGYDADGDGDFDVKVLLDGLGPGGLAKRTAKAKPIKEIK
KRGKAVPSKNEERQKQEDRGKRPDSDTSOKASAAGRDRDKEKASDSS
KSKSWKAVETKAVSSKVAARDKDRGRSSSGHGVSKENGQKRN"
3'UTR 749..1733
polyA_signal 1700..1705
ORIGIN
Query Match 1.3%; Score 31; DB 10; Length 2432;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTTATTAG 333
Db 329 TTTGATGTGGATGATGCCAAAGTTTTATTAG 359

RESULT 53
AC143465/c 158289 bp DNA linear HTG 09-APR-2003
LOCUS Macaca mulatta clone CH250-269N19, *** SEQUENCING IN PROGRESS ***.
DEFINITION AC143465
ACCESSION AC143465
VERSION AC143465.1 GI:29648741
KEYWORDS HTG; HTGS PHASE2; HTGS_PGI.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 158289)
AUTHORS Csuros, M. and Milosavljevic, A.
TITLE Pooled genomic indexing (PGI): mathematical analysis and experiment
design
JOURNAL (in) Guigo, R. and Gusfield, D. (Eds.),
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
REFERENCE 2 (bases 1 to 158289)
AUTHORS Milosavljevic, A., Sodergren, E., Csuros, M., Li, B., Jackson, A.R.,
Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L.,
```

Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, E., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Egan, A., Earnhart, C., Edwards, C.C., Elhaq, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, F., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Masey, E., Metzkler, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Oragunye, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shooshkari, N., Sison, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabot, P., Tamerisa, A., Tameria, K., Tang, H., Tansey, J., Taylor, T., Taylor, I., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Zucheralapati, R., Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
3 (bases 1 to 158289)  
Worley, K.C.

Submitted (09-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center -----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information -----  
Center project name: LBGJ  
Center clone name: CH250-269N19

----- Summary Statistics -----  
Chemistry: Dye-terminator  
Dye-terminator: Big Dye, inf% of reads  
Consensus quality: 7790 bases at least Q40  
Consensus quality: 9211 bases at least Q30  
Consensus quality: 10542 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
\* NOTE: The contigs are based on the application  
of the FGI method using the Human genome (NCBI build 31)  
as the comparative genome.  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces  
is believed to be correct as given, however the sizes  
of the gaps between them are based on estimates that have

\* Provided by the submittor.  
\* This sequence will be replaced  
\* By the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 158289: contig of 158289 bp in length.

## FEATURES

Location/Qualifiers  
1..158289  
/organism="Macaca mulatta"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9544"  
/clone="CH250-269N19" -  
1..158289  
/note="assembly name:CH250-269N19.1B  
CONFIDENCES:\_0.83"

## misc\_feature

## ORIGIN

Query Match 1.3%; Score 30; DB 2; Length 158289;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1310 AGGTCATATGAGAGGTTCCCTGCTTACCCT 1339  
Db 132122 AGGTCATATGAGAGGTTCCCTGCTTACCCT 132093

## RESULT 54

AC095635/c  
LOCUS AC095635 234548 bp DNA linear HTG 09-MAY-2003  
Rattus norvegicus clone CH230-8K1, WORKING DRAFT SEQUENCE, 2  
unordered pieces.

## ACCESSION

AC095635 GI:30467118

HTG: HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

1 (bases 1 to 234548)  
Munzy, D., Marie, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokenkwo, O., Okwuonu, G., Olarnpungsoom, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.L.,

```

Paazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rockey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,D.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,J., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhauser,A., Weiser,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 234548)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 234548)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCWM
Center clone name: CH230-8K1
----- Summary Statistics
Assembly program: Alas
Consensus quality: 222438 bases at least Q40
Consensus quality: 224642 bases at least Q30
Consensus quality: 225972 bases at least Q20
Estimated insert size: 232080; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 2 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 232993: contig of 232993 bp in length
* 232994 233093: gap of unknown length
* 233094 234548: contig of 1455 bp in length.
Location/Qualifiers
1. 234548

```

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-8K1"
1. 1711
/notes="wgs_contig"

misc_feature
ORIGIN

Query Match 1.2%; Score 29; DB 2; Length 234548;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1839 GGTCTTCTTCTGCTGAGTGAACCT 1867
|||||
Db 175547 GGTCTTCTTCTGCTGAGTGAACCT 175519

RESULT 55
AC128056/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-16022, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC128056
AC128056.3 GI-25007693
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus
1 (bases 1 to 275547)
REFERENCE
AUTHORS
Muzny,D.,Marle., Metzker,M., Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyile,M., Cree,A., De Souza,D.,
Davila,M., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duvai,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackeleme,O., Okwono,G., Olarnpungoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,

```

Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1. The Effect of Temperature on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodate	Journal of Chemical Education	1985	Smith, J. D.	2. The Effect of Concentration on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodate	Journal of Chemical Education
3. The Effect of pH on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodate	Journal of Chemical Education	1985	Smith, J. D.	4. The Effect of Catalyst on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodate	Journal of Chemical Education
5. The Effect of Surface Area on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodate	Journal of Chemical Education	1985	Smith, J. D.	6. The Effect of Temperature on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodate	Journal of Chemical Education

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 275547)

REFERENCE	AUTHORS	TITLE	JOURNAL
1	...	...	...
2	...	...	...
3	...	...	...
4	...	...	...
5	...	...	...
6	...	...	...
7	...	...	...
8	...	...	...
9	...	...	...
10	...	...	...
11	...	...	...
12	...	...	...
13	...	...	...
14	...	...	...
15	...	...	...
16	...	...	...
17	...	...	...
18	...	...	...
19	...	...	...
20	...	...	...
21	...	...	...
22	...	...	...
23	...	...	...
24	...	...	...
25	...	...	...
26	...	...	...
27	...	...	...
28	...	...	...
29	...	...	...
30	...	...	...
31	...	...	...
32	...	...	...
33	...	...	...
34	...	...	...
35	...	...	...
36	...	...	...
37	...	...	...
38	...	...	...
39	...	...	...
40	...	...	...
41	...	...	...
42	...	...	...
43	...	...	...
44	...	...	...
45	...	...	...
46	...	...	...
47	...	...	...
48	...	...	...
49	...	...	...
50	...	...	...
51	...	...	...
52	...	...	...
53	...	...	...
54	...	...	...
55	...	...	...
56	...	...	...
57	...	...	...
58	...	...	...
59	...	...	...
60	...	...	...
61	...	...	...
62	...	...	...
63	...	...	...
64	...	...	...
65	...	...	...
66	...	...	...
67	...	...	...
68	...	...	...
69	...	...	...
70	...	...	...
71	...	...	...
72	...	...	...
73	...	...	...
74	...	...	...
75	...	...	...
76	...	...	...
77	...	...	...
78	...	...	...
79	...	...	...
80	...	...	...
81	...	...	...
82	...	...	...
83	...	...	...
84	...	...	...
85	...	...	...
86	...	...	...
87	...	...	...
88	...	...	...
89	...	...	...
90	...	...	...
91	...	...	...
92	...	...	...
93	...	...	...
94	...	...	...
95	...	...	...
96	...	...	...
97	...	...	...
98	...	...	...
99	...	...	...
100	...	...	...

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

**COMMENT**

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>

NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
NOTE: This is a working draft sequence. It currently consists of contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

269305	contig of 269305 bp in length
269306	gap of unknown length
269307	gap of unknown length
269308	contig of 3693 bp in length
273099	gap of unknown length
273199	contig of 1160 bp in length
274359	gap of unknown length
274459	contig of 1089 bp in length

## FEATURES

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-16022"

```

```
misc_feature      1. .1780
                  /note="wgs contig"
misc_feature      266655. .267994
                  /note="wgs contig"
misc_feature      268045. .269305
                  /note="wgs contig"
```

## ORIGIN

Query Match 1.2%; Score 29; DB 2; Length 275547;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 29; Conservative 0. Mismatched 0. Indel 0.

QY 1839 GGTCCTCTCTCTGCTGCCTGAGGATGAAAACCT 1867  
|||  
Db 28865 GGTCCTCTCTCTGCTGCCTGAGGATGAAAACCT 28838

RESULT 56  
C0871665/C

LOCUS	Q0871665	27 bp	DNA	linear	PAT 27-SEP-2004
DEFINITION	Sequence 156 from Patent WO2004078783.				
ACCESSION	Q0871665				
VERSION	Q0871665.1	GI:52745598			

VERSION CQ871665.1 GI:52745698

## KEYWORDS

**SOURCE** synthetic construct

ORGANISM	synthetic construct
<p><i>Escherichia coli</i></p> <p><i>Salmonella enterica</i></p> <p><i>Staphylococcus aureus</i></p> <p><i>Streptococcus pneumoniae</i></p> <p><i>Yersinia enterocolitica</i></p>	<p>100% identity</p> <p>100% identity</p> <p>100% identity</p> <p>100% identity</p> <p>100% identity</p>

other sequences: artificial sequences

## REFERENCE

## ORIGIN

Query Match	1.2%;	Score 27;	DB 6;	Length 27;
Best Local Similarity	100.0%;	Pred. No.	0.018;	

**ay** 286 CTGATGGTGATGGAGATTTTGTATGTGG 312  
|||||  
**bb** 27 CTGATGGTGATGGAGATTTTGTATGTGG 1

RESULT 57  
010058506

Accession	Gene	Size (bp)	Library	Accession
U00000	OCUS	10585	linear	U00000
U00000	EFNITION	13	linear	U00000
U00000	mus musculus aspartyl beta-hydroxylase (ASPH) gene, exon 15.	13	linear	U00000

AF289209  
CCESSION

AF289209.1 GI:11692638

KEYWORDS

EGMENT 5 of 11

**SOURCE** *Mus musculus* (house mouse)

ORGANISM	Mus musculus	Mus musculus
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

Eukaryota: Metazoa: Chordata

Mammalia: Eutheria: Rodentia: Muridae: Murinae: Murini: *Reithrodontomys*: *Reithrodontomys*

## REFERENCE

**AUTHORS** Dinchuk, J. E. Henderson  
(2003-10-2013)

Link J. O'Neil K T Es

FRANK, S.; O'NEILL, K. I.; and HOLLIDAY, C. E. and E. J. 1973.

TIME  
HOLLIS, G.F. and FRIEDMAN  
Accounted for by Friedman?

### III. $\beta$ -hydroxybutyrate

isotom of Asph missing



Friedman, P.A.  
Direct Submission  
Submitted (25-JUL-2000) Applied Biotechnology, DuPont  
Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA

FEATURES  
source  
1. 2013  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
292..378  
/gene="ASPH"  
/number=15

exon

ORIGIN

Query Match 1.2%; Score 27; DB 10; Length 1013;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1077 AAAATTGAGGAGCAGTCAATGCATT 1103  
|||||  
DB 295 AAAATTGAGGAGCAGTCAATGCATT 321  
|||||

RESULT 58  
AC096178  
LOCUS  
DEFINITION Rattus norvegicus clone CH230-11D6, WORKING DRAFT SEQUENCE, 3  
unordered pieces.  
AC096178  
AC096178.6 GI:30522511  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
VERSION  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
1. (bases 1 to 252366)  
Muzny, D., Marie, M., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Altschuler, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinu, H., Divye, K., Egan, A., Escotto, M., Evans, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flag, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoggess, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshehwa, L., Louisedge, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwue, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Soderstrom, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, E., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.  
Direct Submission  
Unpublished  
2. (bases 1 to 252366)  
Worley, K.C.  
Direct Submission  
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3. (bases 1 to 252366)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:24818276.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Project name: GEPP  
Center project name: CH230-11D6  
Center clone name: CH230-11D6  
----- Summary Statistics  
Assembly program: Atlas 3.0  
Consensus quality: 230832 bases at least Q40  
Consensus quality: 233984 bases at least Q30  
Consensus quality: 235460 bases at least Q20  
Estimated insert size: 242518; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 248556: contig of 248556 bp in length  
\* 248557 248556: gap of unknown length  
\* 24857 250663: contig of 2007 bp in length  
\* 250664 250763: gap of unknown length  
\* 250664 252366: contig of 1603 bp in length.  
\* 250764 252366: contig of 1603 bp in length.  
\* 250764 252366: Location/Qualifiers  
1. 252366

FEATURES  
source



```

REFERENCE
AUTHORS Masuda,M. and Igarashi,H.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-1997) Department of Microbiology, Graduate School
of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo
113, Japan

FEATURES
source Location/Qualifiers
1..1519
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="F344"
/db_xref="taxon:10116"
/cell_line="Rat2"
/tissue_type="fibroblast"
/dev_stage="embryo"
1..1519
/gene="PCPE"
CDS 56..1462
/gene="PCPE"
/codon_start=1
/product="procollagen C-proteinase enhancer protein"
/protein_id="AA01592.1"
/db_xref="GI:4102819"
/translation="MLPAALSLGLGPFLLAWLPLARGQTPNYPTRPFLCGGDVTGES
GVYAGEGFNLPPNNKICWITTVPEGVTSLSRFQEDMELHPSGRATSDALEVEVGSGT
SGORLCRGFCGTFRPVAVPGNQVLTMTDEGTGGRGFLLMWSGRATSGETHQFGCG
RNEKAQTTLTNWPESDVPPIGSCSWHIAPSNQVIMLTFGKFDVDPDYCRYDSVS
VNGASVDSSDLKFKFCGDKAPSPISSEGNELLQVFSDLSVTADGFSASYRTLPRDA
VERGAPSCEGAHQGPSRDPKTGTGPKVKPKQVPEKPEGSAPTQATPVAP
DAPSTICPKYKRSGTLQSNFPCSSSLVTGTVKAMVRGPGEGLTVTLVLGVTKTGDL
DIFPASPTSLKFYVPCQMPKKKGASLYLLMGQVENRGFPILPESFVLYRPNDQQ
ILNSLSKRKCPSQPRPDA"

ORIGIN
Query Match 1.0%; Score 24; DB 10; Length 1519;
Best Local Similarity 100.0%; Pred.No. 0.83; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

Qy 36 AGCAGCGGGCAACAGCAGCAGCAGC 59
Db 29 AGCAGCGGCACAGCAGCAGCAGC 6

RESULT 62
MMU19607/c 1680 bp mRNA linear ROD 03-FEB-1996
LOCUS Mus musculus tyrosine kinase Txx (txk) mRNA, complete cds.
DEFINITION
ACCESSION U19607
VERSION U19607.1 GI:643064
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1680)
Sommer,C.L., Huang,K., Shores,E.W., Grinberg,A., Charlick,D.A.,
Kozak,C.A. and Love,P.E.
Murine txk: a protein tyrosine kinase gene regulated by T cell
activation
JOURNAL Oncogene 11 (2), 245-251 (1995)
MEDLINE 95349947
PUBMED 7542761
REFERENCE 2 (bases 1 to 1680)
AUTHORS Sommer,C.L.
DIRECT SUBMISSION
TITLE Submitted (09-JAN-1995) Connie L. Sommers, National Institute of
Child Health and Human Development, National Institutes of Health,
Blodg. 6B, Rm. 2B210, Bethesda, MD 20892, USA
JOURNAL
FEATURES source
Location/Qualifiers
1..1680
/organism="Mus musculus"
/mol_type="mRNA"
/strains="FVB/N"

```

```

/product="Txk"
/protein_id="AAC52264.1"
/db_xref="GI:562125"
/translation="MILSSYSSFQSVLCCCRCSVQKROVRTQISLREBELSEKHS
ORQRPWFAKMGKTQSNRGVQFSRKELPFPQPPDERIQVAKLYDFLPREFGNLA
LKRAEVILIERCDPHWAKRDFNGELIFSNVTENRNLANLEIYWHKNTIRNOT
ERLQKAKAGAFIVDRSHLGSYITISVTRARRHTQSSIKHYQIKNDGQWYITER
HLFSPVELIOYHOYNAAGLISRLRYPIGLLGSCLPATSGFSYEKWEIDPSLAFVKE
IGSQGFVHLGWRAHIVPAIKAINEGMSEEDFIEAAKVMKLSHSLRLVLYGVC
OQKPLYIVTFMENGCLLDYLRERGOQLKALLSMQDICEGMAYLERSCVIHRDLA
ARNCVSSACVVKISDFGMARYVLDDEVYISSGAKFPVKWCPPEVFHFNKYSKSDVM
SFGLYMEVTEGKMPFENKSNLVQVEAISQGRFLYRPHLAPMTIYRVVMYSCWHSPK
GRPTFAELLQVLTEIAETW"
1633..2204

3'UTR
ORIGIN
Query Match 1.0%; Score 24; DB 10; Length 2204;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GAGCAGCGGCAACGACGACGACGAG 58
|||||
Db 110 GAGCAGCGGCAACGACGACGACGAG 87

RESULT 64
MUSRLK/c
LOCUS MUSRLK/c 2221 bp mRNA linear ROD 22-MAY-1995
DEFINITION Mus cookii resting lymphocyte kinase (RLK) gene, complete cds.
ACCESSION L35268
VERSION L35268.1 GI:623442
KEYWORDS protein tyrosine kinase; resting lymphocyte kinase.
SOURCE Mus cookii (Cook's mouse)
ORGANISM Mus cookii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Hu.Q., Davidson,D., Schwartzberg,P.L., Macchiarini,F.,
Lenardo,M.J., Bluestone,J.A. and Matis,L.A.
Identification of Rlk, a novel protein tyrosine kinase with
predominant expression in the T cell lineage
J. Biol. Chem. 270 (4), 1928-1934 (1995)
95130578
PUBMED 7829530
COMMENT On Jan 14, 1995 this sequence version replaced gi:529439.
Original source text: Mus cookii (clone K9A) tetus thymus cDNA to
mRNA.
FEATURES
source
Location/Qualifiers
1..2221
/organism="Mus cookii"
/mol_type="mRNA"
/db_xref="taxon:10098"
/clone="K9A"
/tissue_type="thymus"
/dev_stage="tetus"
1..2221
/gene="RLK"
1..2221
/gene="RLK"
/product="resting lymphocyte kinase"
/notes="predominant expression in T cell lineage; protein
tyrosine kinase"
evidence=experimental
1..48
/gene="RLK"
49..1632
/gene="RLK"
/standard_name="protein tyrosine kinase"
/codon_start=1
/evidence=experimental
/product="resting lymphocyte kinase"
/protein_id="AAA67039.1"
/db_xref="GI:623443"

```

```

/translation="MILSSYSSFQSVLCCCRCSVQKROVRTQISLREBELSEKHS
ORQRPWFAKMGKTQSNRGVQFSRKELPFPQPPDERIQVAKLYDFLPREFGNLA
LKRAEVILIERCDPHWAKRDFNGELIFSNVTENRNLANLEIYWHKNTIRNOT
ERLQKAKAGAFIVDRSHLGSYITISVTRARRHTQSSIKHYQIKNDGQWYITER
HLFSPVELIOYHOYNAAGLISRLRYPIGLLGSCLPATSGFSYEKWEIDPSLAFVKE
IGSQGFVHLGWRAHIVPAIKAINEGMSEEDFIEAAKVMKLSHSLRLVLYGVC
OQKPLYIVTFMENGCLLDYLRERGOQLKALLSMQDICEGMAYLERSCVIHRDLA
ARNCVSSACVVKISDFGMARYVLDDEVYISSGAKFPVKWCPPEVFHFNKYSKSDVM
SFGLYMEVTEGKMPFENKSNLVQVEAISQGRFLYRPHLAPMTIYRVVMYSCWHSPK
GRPTFAELLQVLTEIAETW"
49..1629
/gene="RLK"
/product="resting lymphocyte kinase"
/standard_name="protein tyrosine kinase"
/evidence=experimental
1633..2221
/gene="RLK"
2193..2198
/gene="RLK"
/evidence=experimental
2221
/gene="RLK"
/evidence=experimental

mat_peptide
3'UTR
polyA_signal
polyA_site

ORIGIN
Query Match 1.0%; Score 24; DB 10; Length 2221;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GAGCAGCGGCAACGACGACGACGAG 58
|||||
Db 110 GAGCAGCGGCAACGACGACGACGAG 87

RESULT 65
MUSPTKRL18/c
LOCUS MUSPTKRL18 2342 bp mRNA linear ROD 01-FEB-2003
DEFINITION Mus musculus PTK-RL-18 mRNA for protein tyrosine kinase, complete
cds.
ACCESSION D43963
VERSION D43963.1 GI:604883
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2342)
Higashitsuji,H., Nonoguchi,K., Arii,S., Furutani,M., Kaneko,Y.,
Nakayama,H. and Fujita,J.
Enhanced expression of multiple protein tyrosine kinases in the
mouse regenerating liver : Isolation of PTK-RL-18, a novel
cytoplasmic tyrosine kinase gene of Tec PTK family
Unpublished
2 (bases 1 to 2342)
Higashitsuji,H.
Direct Submission
Submitted (19-DEC-1994) Hiroaki Higashitsuji, Faculty of Medicine,
Kyoto University, Clinical Molecular Biology, Shogoin Kawahara-cho
54, Sakyo-ku, Kyoto, Kyoto 606-8507, Japan
(E-mail:higashitsuji@viru.kyoto-u.ac.jp, Tel:81-75-751-3753,
Fax:81-75-751-3750)
FEATURES
source
Location/Qualifiers
1..2342
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57/BL6"
/db_xref="taxon:10090"
/clone="ptk-18"
/sex="male"
/tissue_type="liver"
/clone_lib="Uni-Zap XR"
/dev_stage="adult"
1..2342
/gene

```

```

CDS
/gene="PTK-RL-18"
168..1751
/gene="PTK-RL-18"
/codon_start=1
/evidence=experimental
/product="protein tyrosine kinase"
/protein_id="BAA07900.1"
/db_xref="GI:604884"
/translation="MISFSDSSFOSVLCCCRCSVKROVRTQISLSEELSEKHS
ORQWPFAKLMGKTQNRGGVQPSKRPPLPQEPDERIQVKAlyDFLPRPGNLA
LRABEYLILERCDFHWKARDRFGNEGLIPSNVTENRNLNLEIYEWYHNITRQI
ERLRQAKGCAFTVRSRLHSGTISVFTARRHTQSSIKHYQIKKNDSGQWYITER
HLFSPVELIQHYNAAGLISRLYPIGLISCLPATSGSYEKWEIDPSLTFVKE
IGSQFGVHLGEWRAHPVAIKAINEGSMSEEDFIEBAKVMKLSHSLVQLYGVCI
QOKPLVYTFEMENCLDYLREKGOLOKALLSMQDDICEGMAYLERSCYTHRDLA
ARNCLVSACVVKISDFGMARYVLDDEYIISSSGAKFPVKWCPPEVHFNKYSKSDVW
SFGVLMEVFTGKMPKNSKNSLQVVAISQGFRLYRPLAPMTIYSVMYSCWHESPK
GRPTFAELLQVLTETATW"
polya_signal
2317..2322
/gene="PTK-RL-18"

ORIGIN
Query Match 1.0%; Score 24; DB 10; Length 2342;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GAGCAGCGGCAACAGCAGCAGCAG 58
|||||
Db 229 GAGCAGCGGCAACAGCAGCAGCAG 206
|||||

RESULT 66
AC013623_2
WPCOMMENT
Sequence split into 5 fragments LOCUS AC013623 Accession AC013623
Fragment Name Begin End
AC013623_0 1 110000
AC013623_1 100001 210000
AC013623_2 200001 310000
AC013623_3 300001 410000
AC013623_4 400001 502225
Continuation (3 of 5) of AC013623 Mus musculus clone RP23-651

Query Match 1.0%; Score 24; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GAGCAGCGGCAACAGCAGCAGCAG 58
|||||
Db 105028 GAGCAGCGGCAACAGCAGCAGCAG 105051
|||||

RESULT 67
AC013623_3
WPCOMMENT
Sequence split into 5 fragments LOCUS AC013623 Accession AC013623
Fragment Name Begin End
AC013623_0 1 110000
AC013623_1 100001 210000
AC013623_2 200001 310000
AC013623_3 300001 410000
AC013623_4 400001 502225
Continuation (4 of 5) of AC013623 Mus musculus clone RP23-651

Query Match 1.0%; Score 24; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GAGCAGCGGCAACAGCAGCAGCAG 58
|||||
Db 5028 GAGCAGCGGCAACAGCAGCAGCAG 5051
|||||
```

```

RESULT 68
CR382132_00
WPCOMMENT
Sequence split into 40 fragments LOCUS CR382132 Accession CR382132
Fragment Name Begin End
CR382132_00 1 110000
CR382132_01 100001 210000
CR382132_02 200001 310000
CR382132_03 300001 410000
CR382132_04 400001 510000
CR382132_05 500001 610000
CR382132_06 600001 710000
CR382132_07 700001 810000
CR382132_08 800001 910000
CR382132_09 900001 1010000
CR382132_10 1000001 1110000
CR382132_11 1100001 1210000
CR382132_12 1200001 1310000
CR382132_13 1300001 1410000
CR382132_14 1400001 1510000
CR382132_15 1500001 1610000
CR382132_16 1600001 1710000
CR382132_17 1700001 1810000
CR382132_18 1800001 1910000
CR382132_19 1900001 2010000
CR382132_20 2000001 2110000
CR382132_21 2100001 2210000
CR382132_22 2200001 2310000
CR382132_23 2300001 2410000
CR382132_24 2400001 2510000
CR382132_25 2500001 2610000
CR382132_26 2600001 2710000
CR382132_27 2700001 2810000
CR382132_28 2800001 2910000
CR382132_29 2900001 3010000
CR382132_30 3000001 3110000
CR382132_31 3100001 3210000
CR382132_32 3200001 3310000
CR382132_33 3300001 3410000
CR382132_34 3400001 3510000
CR382132_35 3500001 3610000
CR382132_36 3600001 3710000
CR382132_37 3700001 3810000
CR382132_38 3800001 3910000
CR382132_39 3900001 4003362
LOCUS CR382132 4003362 bp DNA linear PLN 01-SEP-2004
DEFINITION Yarrowia lipolytica chromosome F of strain CLIB99 of Yarrowia
lipolytica.
ACCESSION CR382132
VERSION CR382132.1 GI:49650673
KEYWORDS genomic DNA.
SOURCE Yarrowia lipolytica CLIB99
ORGANISM Yarrowia lipolytica CLIB99
Saccharomycetales; Dipodascaceae; Yarrowia.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
1 (bases 1 to 4003362)
REFERENCE
AUTHORS
1 Dujon, B., Sherman, D., Fischer, G., Durrans, P., Casaregola, S.,
Lafontaine, I., De Montigny, J., Marck, C., Neuvéglise, C., Talla, E.,
Goffard, N., Frangeul, L., Aigle, M., Anthouard, V., Babour, A.,
Barbe, V., Barnay, S., Blanchin, S., Beckerich, J. M., Beyne, E.,
Bleykasten, C., Boissrame, A., Boyer, J., Cattolico, L.,
Confanier, F., De Daruvar, A., Despons, L., Fabre, E., Fairhead, C.,
Ferry-Dumazet, H., Groppi, A., Hantraye, F., Hennequin, C.,
Jauniaux, N., Joyet, P., Kachouri, R., Kerrest, A., Koszul, R.,
Lemaire, M., Lesur, I., Ma, L., Muller, H., Nicaud, J. M., Nikolski, M.,
Ozias, S., Ozier-Kalogeropoulos, O., Pellenz, S., Potier, S.,
Richard, G. F., Sraub, M. L., Suleau, A., Swennen, D., Tekai, F.,
Wesolowski-Louvel, M., Westhof, E., Wirth, B., Zeniou-Meyer, M.,
Zivanovic, I., Bolotin-Fukuhara, M., Thierry, A., Boucher, C.,
Caudron, B., Scarpelli, C., Gaillardin, C., Weissenbach, J., Wincker, P.
and Souciet J. L.
CONSRMT Genome evolution in yeasts
TITLE
```

JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Nature 430 (6995), 35-44 (2004)  
15229592  
2 (bases 1 to 4003362)  
Genoscope.

Direct Submission  
Submitted (01-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex -FRANCE (E-mail : seqref@genoscope.cns.fr -

Web : www.genoscope.cns.fr)  
This sequence is unfinished. Sequencing was performed at Genoscope  
and annotations were obtained by the genolevures Consortium.

FEATURES  
source

1. 4003362  
/organism="Yarrowia lipolytica CLIB99"  
/mol\_type="genomic DNA"  
/strain="CLIB99"  
/db\_xref="taxon:284591"  
/chromosome="F"

/note="Genoscope sequence ID : YAL10FCHR"

complement (13437..14069)

/locus tag="YAL10F00110g"

complement (13437..14069)

/locus tag="YAL10F00110g"

/note="unnamed protein product; no similarity, possibly  
noncoding, hypothetical start"

/codon\_start=1

/protein\_id="CAG77606.1"

/db\_xref="GI:49650674"

/db\_xref="UniProt/TREMBL:Q6C3F8"

/translation="MTNLFNFKRWLRSGARSAKTKWKKPGVSRSLRFLWTRAT  
CTSPVQEQSPVEILISGPTLLEGDSIPTTLSDPKITEDKKGVNGNSLSQPPQE  
DCVPSVPDTRPKRSCTPLSDLOLRELOSERPDSITLKKRIRKSGQGCD  
TEVLRHPFFAASGADNRSLDCSGIFQCFPEIIEAVSYSLRKNENIPA"

complement (14272..15570)

/locus tag="YAL10F00132g"

complement (14272..15570)

/locus tag="YAL10F00132g"

/note="unnamed protein product; weakly similar to  
tr|Q12161 Saccharomyces cerevisiae YOL054W, hypothetical  
start"

/codon\_start=1

/protein\_id="CAG77607.1"

/db\_xref="GI:49650675"

/db\_xref="UniProt/TREMBL:Q6C3F7"

/translation="MSTDIRFEAHGSDRVAGLLCDVGSLECAICQSIMVPFVLD  
CGHNYCPCLQWVNNSCPCEKTKIRKAPTHSLTLKSLAESITWKMIGVDPISLEK  
PAQRKEQVPAKIDKKRVPFAVSSVHLDEEDGVORCSQHWEEGEQCTN  
CGVWNGSGREAAESINVDDEESDSRQPCEDISYEDGDVDDERVDGDTG  
DVTILSDQASSPAPTSTVQINDSLDSEDEOFLPKABSTEVITLDDSDDEQLQPR  
RKHKMSDEDDHPAVSRKSKSKIFSDSEDDNQOQSSKSKATITFDSEDDERK  
RVSLKXKRDVAVLSDEERESEEDSDYSTDNEGITNGVITGLPFRRRRGDDG  
DENESDDSDSDNDLGGFVVDREADDGSDDEAFYSEY"

complement (15838..18501)

/locus tag="YAL10F00154g"

complement (15838..18501)

/locus tag="YAL10F00154g"

/note="unnamed protein product; similar to sp|Q13396  
Neurospora crassa MSH-2 DNA mismatch repair protein MSH2,  
start by similarity"

/codon\_start=1

/protein\_id="CAG77608.1"

/db\_xref="GI:49650676"

/db\_xref="UniProt/TREMBL:Q6C3F6"

/translation="MSRPFNDKTEERSLTAFAVKGLEEREPDNGGTIRVFERPDG  
YVFGVDARVAVRYRLTALKTTDLKQEVNIIASGQNFIRDALNYGLKVEYIG  
KVQGRWMTSMSPGNLSQVEDLMSGQLTNFVAIAVKTQGTIVGLVLDNNHVLGV  
SEPDNECHNSLESLLIQDVKECITSDKVAIERAGVSRITDAKSWFNANVEEN  
LDNLAKQPTSPSLKNGALSLACLILYLTSDANHGAFITKHTLSQYMKLD  
ASAKALHMPSTKSTSSLYGLNVCKTATGSRTLAQWQKPLMKQIEERHID  
VEFTSDLESIRQLNSTIPDLNLRTRKFMQASLESDVRYVQWATLPHIASGLR  
AAQSELETFITOLDSTITGLQFELVESTIDLSIDSHFEMINPMWESGLNDTKA  
RLACQDKMDKIFASVSDGLGMDKKLFENHHVHGFRITRTDASCGLPERSAVL  
LATLKAGIIFTNELRSLNFTDLSQYKKIQARLAKETIIEIACSYPLERSAVL  
GOLDVLTSPASVAIRNVIRTVVSDSDRCKILITASRHPCLEAQDTFIPNDVHLQDS

KKPLVITGPNMGKSTFIRQGVIVLMNQIGCFVPCDRAEISIFDCILARVAGDSQL  
KGLSTFMEMLSAIKSATDKSLIIDELGRGTSTYDGLGAWAISERHIVQNCFS  
MFATHFELTELAKHPDRVDNLHVAHVGSDDITLLYKVFVGSKSYGTHVAEV  
VFPKTVVNAKRAQELDDVNSGTQGGKYASEDLVAGNKLKELITWKSQIKGDEV  
DGAOLKTTVVVDYKTKQMEQNVFINDALASL"

/locus tag="YAL10F00176g"

complement (19703..20934)

/locus tag="YAL10F00176g"

/note="unnamed protein product; similar to DEHA0G03828g

Debaryomyces Hansenii IPF 785.1 and sp|P20107

Saccharomyces cerevisiae YMR243c ZRC1 zinc- and cadmium

resistance protein, start by similarity"

/codon\_start=1

/protein\_id="CAG77609.1"

/db\_xref="GI:49650677"

/db\_xref="UniProt/TREMBL:Q6C3F5"

/translation="MSLTREIRMIALLIIDIOTFFLEAIVGVHSLALVADSPHML  
NDVFSLIILWAVRAVSRGADSKTYQWAEILGALANAVFLALCLITLLEAIOQ  
LFEPQIITNPKLIAVGTAGLCSTNIVGLLFEHGHAGSHGHDDHDDHARDEEA  
VDTVLASQFTAPTEQTSLLQHPTRSHRSISNDSSEHATFHAKKNEKKVSLNMQ  
GVFLVMDGALGNIGVIATFIMKTYSWKYADPVISLVTIIFSSALPLCRSTS  
SILQAVQPNNAEDNEIVALDGVLEHDLHIWILKEDTFVATLHVAGSDPSEFM  
TSLNDIKKIFHEGINSVTIQPEFNATGTTTPDKHQYHVSVGSLRSANSGLAQ"

complement (19703..20796)

/locus tag="YAL10F00176g"

/number=2

complement (20797..20852)

/locus tag="YAL10F00176g"

/number=1

complement (20853..20934)

/locus tag="YAL10F00176g"

/number=1

/locus tag="YAL10F00198g"

join (22939..22994,23108..25541)

/locus tag="YAL10F00198g"

/note="unnamed protein product; similar to tr|Q08951

Saccharomyces cerevisiae ORF YPL195W Alpha/gamma adaptin,  
start by similarity"

/codon\_start=1

/protein\_id="CAG77610.1"

/db\_xref="GI:49650678"

/db\_xref="UniProt/TREMBL:Q6C3F4"

/translation="MSLTTRDDVRNLRPFVFESTLTDLIKIRAAKDLPAKERFL  
QAVSDCHQEARNDLTVKTLAILKLAILEMGLYDMSWASHVLEVMSPKPFQKRVG  
YLAATOSYRDDDDVLMATNLKRLDTSSTSPVSVLSAIAITVPSLAQDVHPEIV  
KMLSHSKYIRKAVLAMYKI FLYQPEALRTSRLRDLDDSDPSVSVSATVNVICEL  
AKHKSXNYPLAPQYQLLTSTNNMMIKILKLSLAPILPRLPKPLQIMTLIQ  
STGSLQYECINCIQVSGMLGEDDHVAGVCVCSFVNQDQNLKVLGALGI  
KVHPTLVGLQGVILECLENADSTIRERALELANDASEHNVTIVNLLSQOLTH  
ASISYILDMCCRDYSLISDPENFLNKLALVDFLPEDPARIDGLRDLICMRVPM  
REBITGVCYEI FTDGVTYLRNPLVPLVCLSEYPMHQKDLIDFLNSVSHVENS  
GSSNNAVDLAVCVCIKVFQKFCGIIWSSPNELVYKVTWTVIKFSLSYLASDEL  
QDPATQFSELKVI SESLERGVPLLTVELPSFAYELNVPADTKRIIDLELPL  
PELEQTLEVFESDIEDEYVEEFTFPQFASVSGVQKESLIDARKERLEKQKODFFI

Query Match 1.0%; Score 24; DB 8; Length 110000;

Best Local Similarity 100.0%; Pred. No. 0.86;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 AGCAGCGGCAACAGCAGCAGCAGCAGC 59

DB 93631 AGCAGCGGCAACAGCAGCAGCAGCAGC 93654

RESULT 69  
BX663616/c

LOCUS BX663616

DEFINITION Zebrafish DNA sequence from clone DKEY-287H22 in linkage group 7,  
complete sequence.

ACCESSION BX663616

VERSION BX663616.13 GI:47522521

150010 bp DNA linear VPT 19-MAY-2004

AL732455	AL732455.6	GI:21953067
HTG; zinc finger.		
Danio rerio (zebrafish)		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
1 (bases 1 to 192578)		
Skuce, C.		
Direct Submission		
Submitted (20-JUN-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk		
On Jul 24, 2002 this sequence version replaced gi:21912710.		
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.		
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep CH211-225J17 is from a CHORI-211 BAC library		
VECTOR: pTARBAC2.1		
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.		
This sequence is the entire insert of clone CH211-225J17 The true right end of clone CH211-231L18 is at 96893 in this sequence. Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.		
Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml		
Center: Wellcome Trust Sanger Institute		
Center code: SC		
Web site: http://www.sanger.ac.uk		
Contact: zfish-help@sanger.ac.uk		
Location/Qualifiers		
1..192578		
/organism="Danio rerio"		
/mol_type="genomic DNA"		
/db_xref="taxon:7955"		
/clone="CH211-225J17"		
/clone_lib="CHORI-211"		
1..423		
/note="Dr000087 repeat: matches 255..674 of consensus"		
/complement(556..661)		
/notes="drr831 repeat: matches 1..105 of consensus"		
654..1032		
/note="Dr000087 repeat: matches 957..1326 of consensus"		
1033..1089		
/note="19.0 copies 3 mer TAA 11% conserved"		
1114..1127		
/note="2.3 copies 6 mer TTTTAA 28% conserved"		
1145..1207		
/note="CR1-1_DR repeat: matches 4919..4981 of consensus"		
1203..1268		

ACCESSION	AL732455
VERSION	AL732455.6
KEYWORDS	HTG; zinc finger.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
AUTHORS	1 (bases 1 to 192578)
TITLE	Skuce, C.
JOURNAL	Direct Submission
COMMENT	Submitted (20-JUN-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jul 24, 2002 this sequence version replaced gi:21912710. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep CH211-225J17 is from a CHORI-211 BAC library VECTOR: pTARBAC2.1 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. This sequence is the entire insert of clone CH211-225J17 The true right end of clone CH211-231L18 is at 96893 in this sequence. Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted. Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml

FEATURES	Location/Qualifiers
source	1..192578
	/organism="Danio rerio"
	/mol_type="genomic DNA"
	/db_xref="taxon:7955"
	/clone="CH211-225J17"
	/clone_lib="CHORI-211"
repeat_region	1..423
repeat_region	/note="Dr000087 repeat: matches 255..674 of consensus"
repeat_region	/complement(556..661)
repeat_region	/notes="drr831 repeat: matches 1..105 of consensus"
repeat_region	654..1032
repeat_region	/note="Dr000087 repeat: matches 957..1326 of consensus"
repeat_region	1033..1089
repeat_region	/note="19.0 copies 3 mer TAA 11% conserved"
repeat_region	1114..1127
repeat_region	/note="2.3 copies 6 mer TTTTAA 28% conserved"
repeat_region	1145..1207
repeat_region	/note="CR1-1_DR repeat: matches 4919..4981 of consensus"
repeat_region	1203..1268

HTG.	Danio rerio (zebrafish)
Danio rerio	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.	
Auger, K.	
1 (bases 1 to 150010)	
Direct Submission	
Submitted (19-MAY-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk	
On May 19, 2004 this sequence version replaced gi:47109759.	
----- Genome Center	
Center: Wellcome Trust Sanger Institute	
Center code: SC	
Web site: http://www.sanger.ac.uk	
Contact: zfish-help@sanger.ac.uk	
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.	
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.	
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.	
Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-287H22 is from a Zebrafish BAC library	
VECTOR: pindigoBAC-5.	
Location/Qualifiers	
1..150010	
/organism="Danio rerio"	
/mol_type="genomic DNA"	
/db_xref="taxon:7955"	
/clone="DKEY-287H22"	
/clone_lib="DanioKey"	
Query Match	1.0%; Score 24; DB 5; Length 150010;
Best Local Similarity	100.0%; Pred. No. 0.87;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1993 GAAGAGGACAGATCAAAATATTCCA 2016	
Db 29978 GAAGAGGACAGATCAAAATATTCCA 29955	
RESULT 70	
AL732455	192578 bp DNA linear VRT 29-JUL-2003
LOCUS	Zebrafish DNA sequence from clone CH211-225J17 in linkage group 7
DEFINITION	Contains part of a novel gene for a zinc finger protein, complete sequence.

```
repeat_region /note="ACROBAT1 repeat: matches 776. .842 of consensus"
repeat_region complement(1214. .1275)
repeat_region /note="Dr000340 repeat: matches 135. .196 of consensus"
repeat_region complement(1272. .1327)
repeat_region /note="Dr000403 repeat: matches 1. .66 of consensus"
repeat_region complement(1369. .1617)
repeat_region /note="Dr000262 repeat: matches 5. .284 of consensus"
repeat_region 1620. .1631
repeat_region /note="12.0 copies 1 mer T 24% conserved"
repeat_region 1659. .1731
repeat_region /note="Dr000067 repeat: matches 207. .288 of consensus"
repeat_region 1736. .1747
repeat_region /note="2.0 copies 6 mer TTAGTA 24% conserved"
repeat_region 1774. .1795
repeat_region /note="2.8 copies 8 mer TTAAAGA 26% conserved"
repeat_region 1998. .2008
repeat_region /note="2.8 copies 4 mer AGAT 22% conserved"
repeat_region 2017. .2031
repeat_region /note="3.8 copies 4 mer CAGG 30% conserved"
repeat_region 2032. .2102
repeat_region /note="HAT1_DR repeat: matches 575. .647 of consensus"
repeat_region 2103. .2200
repeat_region /note="HAT1_DR repeat: matches 555. .660 of consensus"
repeat_region 2135. .2211
repeat_region /note="Dr000309 repeat: matches 2. .78 of consensus"
repeat_region 2397. .2412
repeat_region /note="8.0 copies 2 mer CT 23% conserved"
repeat_region 2471. .2481
repeat_region /note="2.2 copies 5 mer AATAA 22% conserved"
repeat_region 2930. .2943
repeat_region /note="7.0 copies 2 mer TC 28% conserved"
repeat_region 3063. .3074
repeat_region /note="2.4 copies 5 mer ATATG 24% conserved"
repeat_region 3280. .3407
repeat_region /note="HAT1_DR repeat: matches 521. .659 of consensus"
repeat_region 3371. .3442
repeat_region /note="Dr000309 repeat: matches 2. .69 of consensus"
repeat_region 3408. .3448
repeat_region /note="HAT1_DR repeat: matches 577. .617 of consensus"
repeat_region 3456. .3467
repeat_region /note="2.0 copies 6 mer AACAG 24% conserved"
repeat_region 3629. .3641
repeat_region /note="13.0 copies 1 mer T 26% conserved"
repeat_region 3687. .3696
repeat_region /note="2.5 copies 4 mer ATAT 20% conserved"
repeat_region 3807. .3819
repeat_region /note="2.2 copies 6 mer TGATTC 26% conserved"
repeat_region 3914. .3947
repeat_region /note="8.5 copies 4 mer TAAA 50% conserved"
repeat_region 4157. .4273
repeat_region /note="58.5 copies 2 mer CA 19% conserved"
repeat_region 4174. .4195
repeat_region /note="3.1 copies 7 mer AACTAA 23% conserved"
repeat_region 4274. .4346
repeat_region /note="Dr000236 repeat: matches 358. .433 of consensus"
repeat_region 4642. .4653
repeat_region /note="2.0 copies 6 mer GTTCT 24% conserved"
repeat_region 4726. .4904
repeat_region /note="DNA11TA1_DR repeat: matches 1. .186 of consensus"
repeat_region complement(4960. .5090)
repeat_region /note="ACROBAT1 repeat: matches 751. .876 of consensus"
repeat_region complement(5134. .5259)
repeat_region /note="DNA11TA1_DR repeat: matches 1. .130 of consensus"
repeat_region 5413. .5423
repeat_region /note="2.2 copies 5 mer TACAT 22% conserved"
repeat_region complement(5598. .5681)
repeat_region /note="Dr000009 repeat: matches 770. .857 of consensus"
repeat_region 5803. .5824
repeat_region /note="22.0 copies 1 mer T 44% conserved"
repeat_region 6147. .6167
repeat_region /note="2.3 copies 9 mer ACAGATGG 35% conserved"
repeat_region 6224. .6233
repeat_region /note="2.5 copies 4 mer ACAA 20% conserved"
```

```
repeat_region 6238. .6253
repeat_region /note="16.0 copies 1 mer A 32% conserved"
repeat_region 6626. .6635
repeat_region /note="2.5 copies 4 mer GAAT 20% conserved"
repeat_region 6738. .6760
repeat_region /note="7.7 copies 3 mer TCC 39% conserved"
repeat_region complement(7233. .7249)
repeat_region /note="TDR19 repeat: matches 198. .303 of consensus"
repeat_region 7250. .7336
repeat_region /note="Dr000364 repeat: matches 4. .90 of consensus"
repeat_region 7336. .8045
repeat_region /note="Dr000209 repeat: matches 1. .678 of consensus"
repeat_region 8045. .8138
repeat_region /note="Dr000364 repeat: matches 86. .180 of consensus"
repeat_region 8189. .8203
repeat_region /note="15.0 copies 1 mer T 30% conserved"
repeat_region 8224. .8245
repeat_region /note="5.5 copies 4 mer TCAA 28% conserved"
repeat_region 8507. .8536
repeat_region /note="10.0 copies 3 mer TAC 51% conserved"
repeat_region 8542. .8609
repeat_region /note="22.7 copies 3 mer ATA 118% conserved"
repeat_region 8686. .8703
repeat_region /note="4.5 copies 4 mer CATT 27% conserved"
repeat_region 8705. .8714
repeat_region /note="2.5 copies 4 mer TCGT 20% conserved"
repeat_region 8712. .8769
repeat_region /note="14.5 copies 4 mer TTCC 91% conserved"
repeat_region 8771. .8794
repeat_region /note="6.0 copies 4 mer TTCA 48% conserved"
repeat_region 8828. .8837
repeat_region /note="2.5 copies 4 mer TTAC 20% conserved"
repeat_region 8909. .8919
repeat_region /note="2.8 copies 4 mer TATC 22% conserved"
repeat_region complement(8972. .9131)
repeat_region /note="HEI_DR1 repeat: matches 41. .221 of consensus"
repeat_region 9314. .9344
repeat_region /note="5.2 copies 6 mer ACTATG 30% conserved"
repeat_region 9420. .9432
repeat_region /note="6.5 copies 2 mer AG 26% conserved"
repeat_region 9430. .9441
repeat_region /note="2.4 copies 5 mer AGAA 24% conserved"
repeat_region 9447. .9466
repeat_region /note="2.2 copies 9 mer TTACAGAT 31% conserved"
repeat_region 10208. .10222
repeat_region /note="3.8 copies 4 mer AGGA 21% conserved"
repeat_region 10601. .10611
repeat_region /note="2.8 copies 4 mer CTTC 22% conserved"
repeat_region 10867. .10885
repeat_region /note="3.2 copies 6 mer ACTGTA 29% conserved"
```

Query Match 1.0%; Score 24; DB 5; Length 192578;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1993 GAACAGCAGACATCAATATTCCTCA 2016  
Db 6479 GAACAGCAGACATCAATATTCCTCA 6502

RESULT 71  
AC099698

LOCUS AC099698 222930 bp DNA linear HTG 30-AUG-2003  
DEFINITION Mus musculus chromosome 5 clone RP23-41509 map 5, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 10 unordered pieces.

ACCESSION AC099698  
VERSION AC099698.6 GI:34365951  
KEYWORDS HTG; HTGS PHASE1; HTGS FULLTOP; HTGS\_ACTIVEFIN.  
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 222930)



\* 195077 222930: contig of 27854 bp in length.

FEATURES  
Location/Qualifiers  
1..222930  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="5"  
/map="5"  
/clone="RP23-415O9"  
/clone\_lib="RPCI-23 Female Mouse BAC"

ORIGIN

Query Match 1.0%; Score 24; DB 2; Length 222930;  
Best Local Similarity 100.0%; Pred. No. 0.87; Indels 0; Gaps 0;  
Matches 24; Conservative 0; Mismatches 0

QY 35 GAGCAGCGGCAACAGCACGACGACGAG 58  
|||||  
Db 149599 GAGCAGCGGCAACAGCACGACGACGAG 149622  
|||||

RESULT 72  
AC102441/c

LOCUS  
DEFINITION Mus musculus chromosome 18, clone R224-160112, complete sequence.  
ACCESSION AC102441  
VERSION AC102441.6 GI:32526991  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G., Maclean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Menues,L., Mihova,T., Milenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission  
JOURNAL Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
REFERENCE 3 (bases 1 to 224556)

AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepal,Y., Collamore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Doolley,J.S., Dodge,S., Dooly,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagoian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,

Birren, B., Nusbaum, C. and Lander, E.  
 Mus musculus chromosome 5, clone RP23-41509  
 Unpublished  
 2 (bases 1 to 222930)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,  
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
 Chospel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,  
 Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,  
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
 Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,  
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
 Meneus, L., Mihova, T., Milenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (18-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 222930)  
 Birren, B., Nusbaum, C. and Lander, E.  
 Direct Submission  
 Submitted (30-AUG-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Aug 30, 2003 this sequence version replaced gi:31581768.  
 All repeats were identified using RepeatMasker:  
 Smith, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L17169  
 Center clone name: 415\_0\_9  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 10 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 35963: contig of 35963 bp in length  
 \* 35964 36063: gap of 100 bp  
 \* 36064 40768: contig of 4705 bp in length  
 \* 40769 40868: gap of 100 bp  
 \* 40869 44507: contig of 3639 bp in length  
 \* 44508 44507: gap of 100 bp  
 \* 44508 48988: contig of 5291 bp in length  
 \* 48989 49989: gap of 100 bp  
 \* 49990 54326: contig of 4328 bp in length  
 \* 54327 54426: gap of 100 bp  
 \* 54427 136027: contig of 81601 bp in length  
 \* 136028 136127: gap of 100 bp  
 \* 136128 153799: contig of 17672 bp in length  
 \* 153800 153899: gap of 100 bp  
 \* 153900 191117: contig of 37218 bp in length  
 \* 191118 191217: gap of 100 bp  
 \* 191218 194976: contig of 3759 bp in length  
 \* 194977 195076: gap of 100 bp  
 \* 195077 222930: contig of 27854 bp in length.



```

Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.M.
Direct Submission
Submitted (25-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 25, 2003 this sequence version replaced gi:30018220.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: WBIR
Web code: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11591
Center clone name: 65_I_8
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 28999: contig of 28999 bp in length
* 29000: gap of 100 bp
* 29100 36806: contig of 7707 bp in length
* 36907 36906: gap of 100 bp
* 36907 45591: contig of 8585 bp in length
* 45492 45591: gap of 100 bp
* 45492 48670: contig of 3079 bp in length
* 48670 48670: gap of 100 bp
* 48671 52331: contig of 3561 bp in length
* 48771 52331: gap of 100 bp
* 52332 52332: contig of 15502 bp in length
* 52432 67933: gap of 100 bp
* 67934 68034: contig of 11546 bp in length
* 68034 79579: gap of 100 bp
* 79580 79579: contig of 67278 bp in length
* 79580 146957: gap of 100 bp
* 146958 147057: gap of 100 bp
* 147058 188592: contig of 41535 bp in length
* 188593 207123: gap of 100 bp
* 188593 207123: contig of 18431 bp in length
* 207124 207223: gap of 100 bp
* 207224 289405: contig of 82182 bp in length.
* 207224 Location/Qualifiers
source
1..289405
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP23-65I8"
/clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN
Query Match 1.0%; Score 24; DB 2; Length 289405;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 35 GAGCAGCGGCACAGCAGCAGCAG 58
|||||
Db 193297 GAGCAGCGGCACAGCAGCAGCAG 193274
|||||
RESULT 74
CQ871636
LOCUS CQ871636 23 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 127 from Patent WO2004078783.
ACCESSION CQ871636

```

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CQ871636.1 GI:52745669

synthetic construct  
synthetic construct  
other sequences; artificial sequences.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Murphy, F.J., Sheehan, D.E., Keating, K.E., Hayes, I.E. and Seera, L.E.  
Enzymes involved in apoptosis  
Patent: WO 2004078793-A 127 16-SEP-2004;  
Erix Therapeutics Ltd (IE)

FEATURES  
source

Location/Qualifiers  
1..23  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="QPCR primer"

ORIGIN

Query Match 1.0%; Score 23; DB 6; Length 23;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 CCCGGAGAGAGACAAAGCATGGA 128

Db 1 CCCGGAGAGAGACAAAGCATGGA 23

RESULT 75

G94809/c

LOCUS

S209P6230RC12.T0 C3H/HeJ Mus musculus STS genomic, sequence tagged

DEFINITION

site.

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

Contact: Kerstin Lindblad-Toh

Whitehead Institute for Biomedical Research, Center for Genome

Research

320 Charles Street, Cambridge, MA 02141, USA

Tel: 6172521477

Fax: 6172580903

Email: kersli@genome.wi.mit.edu

Primer A: None

Primer B: None

STS size: 595

Protocol:

WGS-discovery: Paired-end low-coverage whole genome shotgun reads

were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS

reads were placed uniquely on the MGSCv3 C57BL/6J assembly and SNP

detection was carried out by SSAHA-SNP. 225,000 reads were

annotated

as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J

and the strain from which the particular read came. The validation

rate for these SNPs was estimated at approximately 98%.

Location/Qualifiers

1..595

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="C3H/HeJ"

/db\_xref="taxon:10090"

/map="5 22-540 71845760-71846277"

/clone\_lib="C3H/HeJ"

<1..->595

STs  
ORIGIN

Query Match 1.0%; Score 23; DB 11; Length 595;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GAGCAGCGCGCAACAGCAGCAGCA 57

Db 23 GAGCAGCGCGCAACAGCAGCAGCA 1

Search completed: March 25, 2005, 20:35:15  
Job time : 9731 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 16:47:45 ; Search time 6778 Seconds  
(without alignments)  
13051.248 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

- EST.\*
- 1: gb\_est1.\*
  - 2: gb\_est2.\*
  - 3: gb\_est3.\*
  - 4: gb\_est4.\*
  - 5: gb\_est5.\*
  - 6: gb\_est6.\*
  - 7: gb\_est7.\*
  - 8: gb\_est8.\*
  - 9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1803	77.6	2126	9	AY405267 Homo sapi
2	766	33.0	1050	4	BM806632 AGENCOURT
3	689	31.6	1085	5	BM459083 K-EST0070
4	689	29.6	689	4	BM790814 X-EST0070
5	628	27.0	782	7	CK724966 UI-H-ED1
6	603	25.9	918	5	BUI66117 AGENCOURT
7	578	24.9	729	7	CN258120 170005325
8	577	24.8	868	5	BUI75716 AGENCOURT
9	568	24.2	583	5	BP336604 BP336604
10	563	24.2	583	5	BP364957 BP364957
11	556	23.9	859	1	AUI30952 AUI30952
12	543	23.4	651	2	AW996828 QV3-BN004
13	541	23.3	1011	5	EX403636 EX403636
14	538	23.1	589	5	CN258117 170005318
15	537	23.1	589	5	BP266690 BP266690
16	531	22.8	582	5	BP270196 BP270196
17	529	22.8	765	6	CD643948 AGENCOURT
18	517	22.2	517	5	EX479249 EX479249
19	507	21.8	721	4	B1333062 B1333062
20	493	21.2	711	5	CN368832 CN368832
21	491	21.1	868	5	BQ232651 BQ232651
22	476	20.5	520	7	CN368829 CN368829
23	461	19.8	770	7	CN368823 CN368823
24	458	19.7	689	4	BG720220 BG720220

25	455	19.6	455	7	CN274513
26	452	19.4	593	5	BP328261 BP328261
27	450	19.4	857	5	BQ889709 AGENCOURT
28	450	19.4	920	5	BQ722578 AGENCOURT
29	450	19.4	1014	5	BUI90749 AGENCOURT
30	448	19.3	767	1	AUI32666
31	447	19.2	911	5	BUS02606 AGENCOURT
32	438	18.8	519	4	BF962729 QV2-NN004
33	436	18.8	537	2	BE935834 QV2-NN004
34	430	18.5	695	7	CN368827
35	425	18.3	554	1	AL700393 DKF2P686K
36	425	18.3	959	5	BUS02544 AGENCOURT
37	424	18.2	424	1	AL120973 DKF2P762H
38	424	18.2	606	7	CN481470
39	423	18.2	504	4	BF943222 QV2-NN004
40	423	18.2	610	7	CN481475
41	420	18.1	790	1	AUI38660
42	418	18.0	433	5	BQ318605
43	417	17.9	1100	5	BQ278505
44	415	17.9	444	4	BF962946 QV2-NN004
45	409	17.6	570	5	BQ639558
46	402	17.3	740	7	CN368821
47	397	17.1	629	7	CN483754
48	395	17.0	497	5	BQ340475
49	395	17.0	503	5	BQ339559
50	392	16.9	1031	5	BQ223755
51	389	16.7	553	4	BI035330 QV2-NN200
52	385	16.6	409	2	BE084010
53	383	16.5	930	5	BUI838423
54	382	16.4	468	7	CN368831
55	379	16.3	720	6	CA311352
56	370	15.9	626	7	CN368826
57	329	14.2	1048	2	BE876059
58	328	14.1	574	1	AI267311
59	324	13.9	393	7	CN368818
60	322	13.9	580	7	CN481268
61	321	13.8	527	1	AL694159
62	316	13.6	2023	9	AY405268 Pan trogl
63	313	13.5	484	5	BX111251
64	312	13.4	503	7	CF136513
65	311	13.4	682	4	BGI70148 602321860
66	308	13.3	719	7	CRG30295 DKF2P4691
67	308	13.3	737	7	CR559422
68	306	13.2	560	7	CN482623
69	297	12.8	548	7	CN368828
70	295	12.7	509	2	BE018154
71	294	12.7	556	4	BM785890
72	294	12.7	622	1	AV751514
73	294	12.7	862	6	CD512083
74	288	12.4	1010	4	BM462707
75	283	12.2	767	7	CN368820
76	283	12.2	801	5	BUI58234
77	283	12.2	895	4	BI822264
78	278	12.0	652	7	CR769476
79	272	11.7	581	1	AI652752
80	271	11.7	870	5	BX374107
81	270	11.6	468	7	T69872
82	268	11.5	633	7	CN258115
83	265	11.4	715	4	BG701542
84	264	11.4	273	2	BE004012
85	253	10.9	866	4	BG332543
86	252	10.8	774	4	BI193485
87	249	10.7	316	4	BG977347
88	246	10.6	534	6	CB265658
89	244	10.5	614	4	BM785946
90	233	10.0	534	2	BF154495
91	232	10.0	575	2	BF911695
92	232	10.0	644	4	BI546417
93	232	10.0	789	5	BUS43358
94	232	10.0	793	5	BUS42964
95	231	9.9	977	2	CF619299
96	229	9.9	662	2	AW994191
97	226	9.7	769	4	BG166775

```
c 98 225 9.7 404 4 BG979400 BG979400 CM4-CN009
99 222 9.6 549 1 AA176259 2p29h01.r
100 221 9.5 504 6 CD675871 f829b12.y

ALIGNMENTS

RESULT 1
LOCUS AY405267 2126 bp DNA linear GSS 16-DEC-2003
DEFINITION Homo sapiens ASPH gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY405267
VERSION AY405267.1 GI:39761241
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2126)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2126)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submision
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2126
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2126
/gene="ASPH"
/locus_tag="HCM2167"
gene
ORIGIN
Query Match 77.6%; Score 1803; DB 9; Length 2126;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 334 GACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGAGCTGAGCCACACA 393
Db 323 GACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGAGCTGAGCCACACA 382
Qy 394 CTGAGCCCGAGAGAGAGTCTCTGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 453
Db 383 CTGAGCCCGAGAGAGAGTCTCTGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCA 442
Qy 454 AAGAACAAATTCAGTCCCTTCTCCATGAATGGTACAGCCAGAGAGAGAGAGAGAGAGAGAG 513
Db 443 AAGAACAAATTCAGTCCCTTCTCCATGAATGGTACAGCCAGAGAGAGAGAGAGAGAGAGAG 502
Qy 514 ACTTGCAACAGAGAGATGGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
Db 503 ACTTGCAACAGAGAGATGGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 562
Qy 574 TGGGCACTGATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633
Db 563 TGGGCACTGATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 622
```

```
Qy 634 CCGAGCATAGTTTACCACTGGGAGAGACAGTTCACAGACTGTAACTAGGATATGGAAG 693
Db 623 CCGAGCATAGTTTACCACTGGGAGAGACAGTTCACAGACTGTAACTAGGATATGGAAG 682
Qy 694 AGATGATGTCTGAGCAGGAGAAATCCAGATTCAGTGAACCACTAGTAGAGATGAAGAT 753
Db 683 AGATGATGTCTGAGCAGGAGAAATCCAGATTCAGTGAACCACTAGTAGAGATGAAGAT 742
Qy 754 TGCACCATGATACAGATGATGTAAACATACCACTCTATGAGGAAACAGCAGATATATGAAC 813
Db 743 TGCACCATGATACAGATGATGTAAACATACCACTCTATGAGGAAACAGCAGATATATGAAC 802
Qy 814 CTCTAGAAATGAAGGGATAGAAATACAGAACTAATCTGCTCCCTCCTGAGGATATCTCTG 873
Db 803 CTCTAGAAATGAAGGGATAGAAATACAGAACTAATCTGCTCCCTCCTGAGGATATCTCTG 862
Qy 874 TAGAAGATTCACAGGTAATTTGTAGAGAGAGTAGCAATTTTCTGCTGGAGAGACAGCAGG 933
Db 863 TAGAAGATTCACAGGTAATTTGTAGAGAGAGTAGCAATTTTCTGCTGGAGAGACAGCAGG 922
Qy 934 AAGTACCACAGAGAAACAAATAGAAAAACAGATCCAGAAACAAAAAGCAAAAGTTAAGA 993
Db 923 AAGTACCACAGAGAAACAAATAGAAAAACAGATCCAGAAACAAAAAGCAAAAGTTAAGA 982
Qy 994 AAAAGAGCCTAACTTTTAAATTAATTTGATGAAGACTATTAAGCTGAACCTTGATGCTG 1053
Db 983 AAAAGAGCCTAACTTTTAAATTAATTTGATGAAGACTATTAAGCTGAACCTTGATGCTG 1042
Qy 1054 CAGAAAACTCCGTAAAGGGGAAATTTGAGGAAGCAGTGAATGATTTTAAAGAACTAG 1113
Db 1043 CAGAAAACTCCGTAAAGGGGAAATTTGAGGAAGCAGTGAATGATTTTAAAGAACTAG 1102
Qy 1114 TAGCGAAATACCTTCAGAGTCCACGAGCAAGATATGGGAGGCGCAGTGTGAGGATGATT 1173
Db 1103 TAGCGAAATACCTTCAGAGTCCACGAGCAAGATATGGGAGGCGCAGTGTGAGGATGATT 1162
Qy 1174 TGCTGAGAGAGAGAGAGATGATGAGGTGCTAGCTGAGGCCATCGAGACCTACCAAGAGG 1233
Db 1163 TGCTGAGAGAGAGAGAGATGATGAGGTGCTAGCTGAGGCCATCGAGACCTACCAAGAGG 1222
Qy 1234 TGCCAGCCTACCTGATGCTCCCTGAGAGCCTGCTGAGAGCTGAGTGTGAACGCTCGCTCAG 1293
Db 1223 TGCCAGCCTACCTGATGCTCCCTGAGAGCCTGCTGAGAGCTGAGTGTGAACGCTCGCTCAG 1282
Qy 1294 ACAGGCAACAATTTCTAGGTTCATATGAGAGTTTCCCTGCTTACCTGCGAGAGATTAGTTC 1353
Db 1283 ACAGGCAACAATTTCTAGGTTCATATGAGAGTTTCCCTGCTTACCTGCGAGAGATTAGTTC 1342
Qy 1354 AACTATTTCCCAATGATCTTCTTAAATAATGACCTTGGCGTGGGATACCTCTTGATAG 1413
Db 1343 AACTATTTCCCAATGATCTTCTTAAATAATGACCTTGGCGTGGGATACCTCTTGATAG 1402
Qy 1414 GAGATATGACATGCAAGAGAAAGTTTATGAAGAGGTGCTGAGTGTGACACCTAATGATG 1473
Db 1403 GAGATATGACATGCAAGAGAAAGTTTATGAAGAGGTGCTGAGTGTGACACCTAATGATG 1462
Qy 1474 GCTTTGCTAAAGTCCATTTATGGCTTCATCTCTGAAGGCACAGAAACAAATTTCTGAGAGCA 1533
Db 1463 GCTTTGCTAAAGTCCATTTATGGCTTCATCTCTGAAGGCACAGAAACAAATTTCTGAGAGCA 1522
Qy 1534 TCCCATATTTAAAGGAAGGATAGATCCGAGATCTCGGACCTGATGATGGGAGATTTT 1593
Db 1523 TCCCATATTTAAAGGAAGGATAGATCCGAGATCTCGGACCTGATGATGGGAGATTTT 1582
Qy 1594 ATTTCCACCTGGGGATGCCATGCGAGGGTTGGGAAACAAAGAGGCAATATAAGTGGTATG 1653
Db 1583 ATTTCCACCTGGGGATGCCATGCGAGGGTTGGGAAACAAAGAGGCAATATAAGTGGTATG 1642
Qy 1654 AGCTTGGGCAACAGAGAGGACATTTTGCATCTGTCTGGCAACCGCTCACTCTACATGTGA 1713
Db 1643 AGCTTGGGCAACAGAGAGGACATTTTGCATCTGTCTGGCAACCGCTCACTCTACATGTGA 1702
Qy 1714 ATGGACTGAAAGCACACCTTTGGTGGACCCCAAGAAACCGGCTACACAGAGTTAGTAA 1773
```

```
Db 1703 ATGGACTGAAGACACAGCCTTGGTGGACCCCAAAAGAAACGGGCTACACAGAGTTAGTAA 1762
Qy 1774 AGTCTTTAGAAAGAACTGGAAAGTTAATCCGAGATGAAGGCCCTTGCAGTGATGATGAAG 1833
Db 1763 AGTCTTTAGAAAGAACTGGAAAGTTAATCCGAGATGAAGGCCCTTGCAGTGATGATGAAG 1822
Qy 1834 CCAAGGCTCTCTCTGCTGAGGATGAAGAACTGAGGAAAGGGAAGGGAAGGGAAGGGAAGG 1893
Db 1823 CCAAGGCTCTCTCTGCTGAGGATGAAGAACTGAGGAAAGGGAAGGGAAGGGAAGGGAAGG 1882
Qy 1894 TCACGCTGGGAGCAGAGGAAGAAATGAAATGCCCTGCAAGAGGAGCTCCTAAAACCT 1953
Db 1883 TCACGCTGGGAGCAGAGGAAGAAATGAAATGCCCTGCAAGAGGAGCTCCTAAAACCT 1942
Qy 1954 GTACCTTACTAGAAAGTTCCCGGAGACACACAGGATGCGAAGAGGAGCAGATCAAAATATT 2013
Db 1943 GTACCTTACTAGAAAGTTCCCGGAGACACACAGGATGCGAAGAGGAGCAGATCAAAATATT 2002
Qy 2014 CCATCATGACCCCGGAGCTCAGTGTCGCCGACACACAGGCGCCACAACTGCAGGCTCC 2073
Db 2003 CCATCATGACCCCGGAGCTCAGTGTCGCCGACACACAGGCGCCACAACTGCAGGCTCC 2062
Qy 2074 GAATGACCTGGGCTGGTGTATCCCAAGGAAGGTCGAAGATTCGATGTCGAACGAGA 2133
Db 2063 GAATGACCTGGGCTGGTGTATCCCAAGGAAGGTCGAAGATTCGATGTCGAACGAGA 2122
Qy 2134 CCA 2136
Db 2123 CCA 2125
```

## RESULT 2

BM806632  
LOCUS  
DEFINITION AGENCOURT\_6543030 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5548943  
5' mRNA sequence.

ACCESSION BM806632.1 GI:19123455

## KEYWORDS

SOURCE EST.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

## FEATURES

Source: ATCC/DCTD/DTDP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12258 row: k column: 24  
High quality sequence stop: 649.  
Location/Qualifiers  
1..1050  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5548943"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;  
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

## ORIGIN

Query Match 33.0%; Score 766; DB 4; Length 1050;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 766; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 865 ATATTCCTGTAGAGATTCACAGGTAATTTGTAGAAAGTAAGCATTTTTCTGTGGAAG 924  
Db 1 ATATTCCTGTAGAGATTCACAGGTAATTTGTAGAAAGTAAGCATTTTTCTGTGGAAG 60  
Qy 925 AACAGCAGGAGTACCAACAGAAACAAATAGAAAAACAGATGATCCAGAAACAAAAAGCAA 984  
Db 61 AACAGCAGGAGTACCAACAGAAACAAATAGAAAAACAGATGATCCAGAAACAAAAAGCAA 120.  
Qy 985 AAGTTAAGAAAAAGAGAGCCCTTAACCTTTTAAATTAATTTGTAGAGCTATTAAAGCTGAAC 1044  
Db 121 AAGTTAAGAAAAAGAGAGCCCTTAACCTTTTAAATTAATTTGTAGAGCTATTAAAGCTGAAC 180  
Qy 1045 TTGATGCTGCAGAAAAAAGCTCCGTAAAGGGGAAAAATTTGAGGAAGCAAGTGAATGCAATTTA 1104  
Db 181 TTGATGCTGCAGAAAAAAGCTCCGTAAAGGGGAAAAATTTGAGGAAGCAAGTGAATGCAATTTA 240  
Qy 1105 AAGAACTAGTAGCAAAATACCTTCAGAGTCCACGAGCAAGATATGGGAAGGGCGCAGTGTG 1164  
Db 241 AAGAACTAGTAGCAAAATACCTTCAGAGTCCACGAGCAAGATATGGGAAGGGCGCAGTGTG 300  
Qy 1165 AGGATGATTTGGCTGAGAAAGAGAGAAAGTAAATGAGGTGCTACGTGGAGCCATCGAGACCT 1224  
Db 301 AGGATGATTTGGCTGAGAAAGAGAGAAAGTAAATGAGGTGCTACGTGGAGCCATCGAGACCT 360  
Qy 1225 ACCAAGAGTGGCAGGCTACCTGATGCTCCCTGCAGACCTGCTGAAGCTGAGTTGAAGC 1284  
Db 361 ACCAAGAGTGGCAGGCTACCTGATGCTCCCTGCAGACCTGCTGAAGCTGAGTTGAAGC 420  
Qy 1285 GTGCTCAGACAGGCAAAATTTCTAGGTCAATATGAGAGGTTCCTGCTTACCCTGCAGA 1344  
Db 421 GTGCTCAGACAGGCAAAATTTCTAGGTCAATATGAGAGGTTCCTGCTTACCCTGCAGA 480  
Qy 1345 GATTAGTTCAACTATTTCCTCAATGATCTTCCTTTAAAAAATGACCTTTGGCGTGGGATACC 1404  
Db 481 GATTAGTTCAACTATTTCCTCAATGATCTTCCTTTAAAAAATGACCTTTGGCGTGGGATACC 540  
Qy 1405 TCTTGATAGAGATTAATGACAAATGCAAGAAAGTTTATGAGAGGTGCTGAGTGTGACAC 1464  
Db 541 TCTTGATAGAGATTAATGACAAATGCAAGAAAGTTTATGAGAGGTGCTGAGTGTGACAC 600  
Qy 1465 CTATGATGGCTTTGCTAAAGTCCATTTATGGCTTCATCTCTGAAGGCAAGAAACAAATTTG 1524  
Db 601 CTATGATGGCTTTGCTAAAGTCCATTTATGGCTTCATCTCTGAAGGCAAGAAACAAATTTG 660  
Qy 1525 CTGAGAGCATCCCATATTATTAAGGAAGGAATAGAAATCCGGAGATCTCGGCACTGATGATG 1584  
Db 661 CTGAGAGCATCCCATATTATTAAGGAAGGAATAGAAATCCGGAGATCTCGGCACTGATGATG 720  
Qy 1585 GGAGATTTTATTTCCACCTGGGGGATGCGCATGAGAGGGTTGGGA 1630  
Db 721 GGAGATTTTATTTCCACCTGGGGGATGCGCATGAGAGGGTTGGGA 766

## RESULT 3

## BX459083

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

BX459083 1085 bp mRNA linear EST 06-MAY-2004  
BX459083 Homo sapiens PLACENTA Homo sapiens cDNA clone CS05E012YB13  
5-PRIME, mRNA sequence.

ACCESSION BX459083

VERSION BX459083.2 GI:47064816

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1085)

AUTHORS Li, W.B.; Gruber, C.; Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization



Unpublished (2001)  
On May 22, 2003 this sequence version replaced gi:31027087.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 4537.1  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna/s=CS0DE012CA07QP1&c=4537.r.  
FEATURES  
Location/Qualifiers  
1..1085  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE012B13"  
/tissue\_type="PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."  
ORIGIN  
Query Match 31.6%; Score 734; DB 5; Length 1085;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 784; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 162 GGAACCTTCATCTTCACGTGGTTATGGTATGCTGCGGCTTGACATCTGTA 221  
DB 50 GGAACCTTCATCTTCACGTGGTTATGGTATGCTGCGGCTTGACATCTGTA 109  
QY 222 GCTGCTGTTGTTGATCTTCTGCTATGAGGAGTTCTAGGAAACTAGGAATCTAT 281  
DB 110 GCTGCTGTTGTTGATCTTCTGCTATGAGGAGTTCTAGGAAACTAGGAATCTAT 169  
QY 282 GATGCTGATGTCATGAGATTTTATGTTGATGTCGCAAAAGTTTATTAGGACTTAA 341  
DB 170 GATGCTGATGTCATGAGATTTTATGTTGATGTCGCAAAAGTTTATTAGGACTTAA 229  
QY 342 GAGAGATCTATTTAGAGCCAGTCGCCGACAGAGAGGCTGAGCCACACTGAGCCC 401  
DB 230 GAGAGATCTATTTAGAGCCAGTCGCCGACAGAGAGGCTGAGCCACACTGAGCCC 289  
QY 402 GAGGACAGGTTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461  
DB 290 GAGGACAGGTTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 349  
QY 462 ATTCACTGCTCTCTCCATGAAATGTTACACGAGAACATGTTGAGGAGAGAGACTTGA 521  
DB 350 ATTCACTGCTCTCTCCATGAAATGTTACACGAGAACATGTTGAGGAGAGAGACTTGA 409  
QY 522 CAAGAAGATGAGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACT 581  
DB 410 CAAGAAGATGAGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACT 469  
QY 582 GATGTTAGATGATGATTTGAGACCTTGAACCTTGAAGATATCTCATGAAGAAACCGAGCAT 641  
DB 470 GATGTTAGATGATGATTTGAGACCTTGAACCTTGAAGATATCTCATGAAGAAACCGAGCAT 529  
QY 642 AGTTACACCTGGAAGAGAGAGAGTTTCAAGAGCTGTAATCAGGATATGAGAGATGATG 701  
DB 530 AGTTACACCTGGAAGAGAGAGAGTTTCAAGAGCTGTAATCAGGATATGAGAGATGATG 589  
QY 702 TCTGAGCAGGAAATCCAGATTCAGTGACACCTAGTAGAGATGAAAGATTTGACCAT 761  
DB 590 TCTGAGCAGGAAATCCAGATTCAGTGACACCTAGTAGAGATGAAAGATTTGACCAT 649

JOURNAL  
COMMENT  
Unpublished (2001)  
On May 22, 2003 this sequence version replaced gi:31027087.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 4537.1  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna/s=CS0DE012CA07QP1&c=4537.r.  
FEATURES  
Location/Qualifiers  
1..1085  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE012B13"  
/tissue\_type="PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."  
ORIGIN  
Query Match 31.6%; Score 734; DB 5; Length 1085;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 784; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 162 GGAACCTTCATCTTCACGTGGTTATGGTATGCTGCGGCTTGACATCTGTA 221  
DB 50 GGAACCTTCATCTTCACGTGGTTATGGTATGCTGCGGCTTGACATCTGTA 109  
QY 222 GCTGCTGTTGTTGATCTTCTGCTATGAGGAGTTCTAGGAAACTAGGAATCTAT 281  
DB 110 GCTGCTGTTGTTGATCTTCTGCTATGAGGAGTTCTAGGAAACTAGGAATCTAT 169  
QY 282 GATGCTGATGTCATGAGATTTTATGTTGATGTCGCAAAAGTTTATTAGGACTTAA 341  
DB 170 GATGCTGATGTCATGAGATTTTATGTTGATGTCGCAAAAGTTTATTAGGACTTAA 229  
QY 342 GAGAGATCTATTTAGAGCCAGTCGCCGACAGAGAGGCTGAGCCACACTGAGCCC 401  
DB 230 GAGAGATCTATTTAGAGCCAGTCGCCGACAGAGAGGCTGAGCCACACTGAGCCC 289  
QY 402 GAGGACAGGTTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461  
DB 290 GAGGACAGGTTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 349  
QY 462 ATTCACTGCTCTCTCCATGAAATGTTACACGAGAACATGTTGAGGAGAGAGACTTGA 521  
DB 350 ATTCACTGCTCTCTCCATGAAATGTTACACGAGAACATGTTGAGGAGAGAGACTTGA 409  
QY 522 CAAGAAGATGAGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACT 581  
DB 410 CAAGAAGATGAGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACT 469  
QY 582 GATGTTAGATGATGATTTGAGACCTTGAACCTTGAAGATATCTCATGAAGAAACCGAGCAT 641  
DB 470 GATGTTAGATGATGATTTGAGACCTTGAACCTTGAAGATATCTCATGAAGAAACCGAGCAT 529  
QY 642 AGTTACACCTGGAAGAGAGAGAGTTTCAAGAGCTGTAATCAGGATATGAGAGATGATG 701  
DB 530 AGTTACACCTGGAAGAGAGAGAGTTTCAAGAGCTGTAATCAGGATATGAGAGATGATG 589  
QY 702 TCTGAGCAGGAAATCCAGATTCAGTGACACCTAGTAGAGATGAAAGATTTGACCAT 761  
DB 590 TCTGAGCAGGAAATCCAGATTCAGTGACACCTAGTAGAGATGAAAGATTTGACCAT 649

ORIGIN

```
Query Match      29.6%; Score 689; DB 4; Length 689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1219 AGACCTACCAAGAGTGGCCAGCTACCTGATGTCCTCCGACAGCTGCTGAAGTGAAT 1278
Db 1 AGACCTACCAAGAGTGGCCAGCTACCTGATGTCCTCCGACAGCTGCTGAAGTGAAT 60
QY 1279 TGAAGGTCGCTCAGACAGGCAACAATTTCTAGGTCATATGAGAGGTTCCCTTACCC 1338
Db TGAAGGTCGCTCAGACAGGCAACAATTTCTAGGTCATATGAGAGGTTCCCTTACCC 120
QY 1339 TCCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTTAAAAAATGACCTTGGCGTG 1398
Db TCCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTTAAAAAATGACCTTGGCGTG 180
QY 1399 GATACCTCTCATAGGAGATAATGACATGCAAGAAAGTTTATGAAGAGTCTGAGTG 1458
Db GATACCTCTCATAGGAGATAATGACATGCAAGAAAGTTTATGAAGAGTCTGAGTG 240
QY 1459 TCACACCTAATGATGCTTTGCTAAAGTCCATTTATGCTTCCATGCTGAGGCAAGCA 1518
Db TCACACCTAATGATGCTTTGCTAAAGTCCATTTATGCTTCCATGCTGAGGCAAGCA 300
QY 1519 AAATGCTGAGAGATCCCATATTTAAAGAAAGGAATAGAAATCCGGAGATCCTGGCACTG 1578
Db AAATGCTGAGAGATCCCATATTTAAAGAAAGGAATAGAAATCCGGAGATCCTGGCACTG 360
QY 1579 ATGATGGAGATTTTATTTCCACCTGGGGATGCCATGACAGAGGTTGGGAAACAAAGAGG 1638
Db ATGATGGAGATTTTATTTCCACCTGGGGATGCCATGACAGAGGTTGGGAAACAAAGAGG 420
QY 1639 CATATAGTGTATGAGCTGGGCAACAGAGAGCACTTTGATCTGTCTGCAACGCT 1698
Db CATATAGTGTATGAGCTGGGCAACAGAGAGCACTTTGATCTGTCTGCAACGCT 480
QY 1699 CACTCTACATGTGATGAGCTGAAAGCAAGCCTTGGTGAGCCCAAGAAAGAGGCT 1758
Db CACTCTACATGTGATGAGCTGAAAGCAAGCCTTGGTGAGCCCAAGAAAGAGGCT 540
QY 1759 ACAGAGTGTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAAGGCTTG 1818
Db ACAGAGTGTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAAGGCTTG 600
QY 1819 CAGTGTAGGATAAAGCAAGGCTCTTCTGCTGAGGATGAACCTGAGGGAAG 1878
Db CAGTGTAGGATAAAGCAAGGCTCTTCTGCTGAGGATGAACCTGAGGGAAG 660
QY 1879 GGGACTGGAGCCAGTTCAGCTGTGGCAG 1907
Db GGGACTGGAGCCAGTTCAGCTGTGGCAG 689
```

```
RESULT 5
CK724966/c
LOCUS
DEFINITION
  CK724966
ACCESSION
  CK724966
VERSION
  CK724966.1
KEYWORDS
  EST,
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 782)
  NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Daniela S. Gerhard, Ph.D.
  Office of Cancer Genomics
```

National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/cgap.html>  
Seq primer: M13 FORWARD  
POLYA=Yes.

#### FEATURES

##### Location/Qualifiers

```
1..782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-ED1-axp-c-22-0-UI"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED1"
/notes="Organ: Left Pubic Bone; Vector: pT7T3-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library
containing the following tissue(s): Chondrosarcoma cell
line C95. The library was constructed according to
Bonardo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GCTCAGGCT.
TAG TISSUE=chondrosarcoma
TAG LIB=UI-H-ED1
TAG_SEQ=CGTCAAGGCT"
```

#### ORIGIN

```
Query Match      27.0%; Score 628; DB 7; Length 782;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 678; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1233 GTGCCAGCTACCTGATGCTCCTCGACACCTGCTGAAGCTGATTTGAAGCTGCTCA 1292
Db 694 GTGCCAGCTACCTGATGCTCCTCGACACCTGCTGAAGCTGATTTGAAGCTGCTCA 635
QY 1293 GACAGGCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCTCGCAGAGATTAGTT 1352
Db 634 GACAGGCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCTCGCAGAGATTAGTT 575
QY 1353 CAACATTTCCCAATGATACTTCCCTTAAAAAATGACCTTGGCGTGGGATACCTCTTGATA 1412
Db 574 CAACATTTCCCAATGATACTTCCCTTAAAAAATGACCTTGGCGTGGGATACCTCTTGATA 515
QY 1413 GGAGATAATGACATGCAAGAAAGGTTTATGAAGAGGCTGAGTCTGACACTAATGAT 1472
Db 514 GGAGATAATGACATGCAAGAAAGGTTTATGAAGAGGCTGAGTCTGACACTAATGAT 455
QY 1473 GGCTTTGCTTAAAGTCCATTTATGGCTTCATCTTGAAGGCAAGAAATTCCTGAGAGC 1532
Db 454 GGCTTTGCTTAAAGTCCATTTATGGCTTCATCTTGAAGGCAAGAAATTCCTGAGAGC 395
QY 1533 ATCCCATATTTAAAGAAAGGATAGAAATCCGAGATCCTGCGACTGATGATGGGAGATT 1592
Db 394 ATCCCATATTTAAAGAAAGGATAGAAATCCGAGATCCTGCGACTGATGATGGGAGATT 335
QY 1593 TATTTCCACTCGGGGATGCCATGACAGAGGTTGGGAAACAAAGAGGCATATAAGTGAT 1652
Db 334 TATTTCCACTCGGGGATGCCATGACAGAGGTTGGGAAACAAAGAGGCATATAAGTGAT 275
QY 1653 GAGCTGGGGCAAGAGAGGACACTTTCATCTGCTGGCAACGCTCACTCTACAAATGTG 1712
```



from undifferentiated hES cell lines H1 (p32), H7 (p29),  
and H9 (p26) maintained in feeder-free conditions"

## ORIGIN

```
Query Match      24.9%; Score 578; DB 7; Length 729;
Best Local Similarity 99.8%; Pred. No. 1.2e-298;
Matches 628; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1696 GCTCACTCTACAACTGGAATGGAACAGCAGCCTTGGTGGACCCCAAAAGAAACGG 1755
Db      |||
QY 3 GCTCACTCTACAACTGGAATGGAACAGCAGCCTTGGTGGACCCCAAAAGAAACGG 62
Db      |||
QY 1756 GCTACACAGAGTTAGTAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAAGGCC 1815
Db      |||
QY 63 GCTACACAGAGTTAGTAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAAGGCC 122
Db      |||
QY 1816 TTGCAGTGATGATAAAGCCAAAGGCTCTTCTGCTGCTGAGGATGAAACCTGAGGGAAA 1875
Db      |||
QY 123 TTGCAGTGATGATAAAGCCAAAGGCTCTTCTGCTGCTGAGGATGAAACCTGAGGGAAA 182
Db      |||
QY 1876 AAGGGGACTGGAGCCAGTTCAAGCTGTGGCAGCAGGAGAGAAATGAAATGCCTGCA 1935
Db      |||
QY 183 AAGGGGACTGGAGCCAGTTCAAGCTGTGGCAGCAGGAGAGAAATGAAATGCCTGCA 242
Db      |||
QY 1936 AAGGAGCTCTTAAACCTGTACTTACTAGAAAGTTCCCGAGACACAGGATGAGAA 1995
Db      |||
QY 243 AAGGAGCTCTTAAACCTGTACTTACTAGAAAGTTCCCGAGACACAGGATGAGAA 302
Db      |||
QY 1996 GAGGACAGATCAAAATATTCATCATGACACCCGGGACTCAGCTGTGGCCGACACAGGGC 2055
Db      |||
QY 303 GAGGACAGATCAAAATATTCATCATGACACCCGGGACTCAGCTGTGGCCGACACAGGGC 362
Db      |||
QY 2056 CCACAAATCGCAGGCTCCGAATGCACTGGCTTGGTGTATTCACAGAGAGGCTGCA 2115
Db      |||
QY 363 CCACAAATCGCAGGCTCCGAATGCACTGGCTTGGTGTATTCACAGAGAGGCTGCA 422
Db      |||
QY 2116 TTGCATGTGCCAAGCAGCAGGAGCTGGAGAGGAGGAGGCTCATCTTTGATGACT 2175
Db      |||
QY 423 TTGCATGTGCCAAGCAGCAGGAGCTGGAGAGGAGGAGGCTCATCTTTGATGACT 482
Db      |||
QY 2176 CTTTGACGACAGGATATGCGAGGATGCTCATCTTTCCGGCTGATATTCATGCTGGATG 2235
Db      |||
QY 483 CTTTGACGACAGGATATGCGAGGATGCTCATCTTTCCGGCTGATATTCATGCTGGATG 542
Db      |||
QY 2236 TTGGCATCCGGAATGACACACAGCAGAGAGCAGGCTTCCAGCAATTTAGCATGAAT 2295
Db      |||
QY 543 TTGGCATCCGGAATGACACACAGCAGAGAGCAGGCTTCCAGCAATTTAGCATGAAT 602
Db      |||
QY 2296 TCATGCAAGCTTGGGAACTCTGGAGAGA 2324
Db      |||
QY 603 TCATGCAAGCTTGGGAACTCTGGAGAGA 631
Db      |||
```

## RESULT 8

LOCUS BUI75716  
DEFINITION AGENCOURT\_7953975 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6010847  
5', mRNA sequence.

ACCESSION BUI75716

VERSION BUI75716.1

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 868)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: DCTD/PTG/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13199 row: i column: 24  
High quality sequence stop: 547.  
Location/Qualifiers  
1. 868  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6010847"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_68"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.8 kb. Library constructed by Life  
Technologies."

## ORIGIN

```
Query Match      24.8%; Score 577; DB 5; Length 868;
Best Local Similarity 99.8%; Pred. No. 4.4e-298;
Matches 627; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1692 CAACGCTCACTCTACAACTGGAATGGAACAGCAGCCTTGGTGGACCCCAAAAGAA 1751
Db      |||
QY 1752 ACGGGCTACACAGAGTTAGTAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAA 1811
Db      |||
QY 61 ACGGGCTACACAGAGTTAGTAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAA 120
Db      |||
QY 1812 GGCCTTCAGTGATGGATTAAGCCAAAGGCTCTTCTGCTGCTGAGGATGAAACCTTGAGG 1871
Db      |||
QY 121 GGCCTTCAGTGATGGATTAAGCCAAAGGCTCTTCTGCTGCTGAGGATGAAACCTTGAGG 180
Db      |||
QY 1872 GAAAAGGGGACTTGGAGCCAGTTTACGCTGTGGCAGCAGGAGAGAGAAATGAAATGCC 1931
Db      |||
QY 181 GAAAAGGGGACTTGGAGCCAGTTTACGCTGTGGCAGCAGGAGAGAGAAATGAAATGCC 240
Db      |||
QY 1932 TGCAAAAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGAGACAAAGATGC 1991
Db      |||
QY 241 TGCAAAAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGAGACAAAGATGC 300
Db      |||
QY 1992 AGAAGAGGACAGATCAAAATATTCATCATGCAACCCGGGACTCAGCTGTGGCCGACACA 2051
Db      |||
QY 301 AGAAGAGGACAGATCAAAATATTCATCATGCAACCCGGGACTCAGCTGTGGCCGACACA 360
Db      |||
QY 2052 GGGCCACAAAATTCGAGGCTCCGAATGCACTGGGCTTGGTGTATTCCTCAAGGAAGGCTGC 2111
Db      |||
QY 361 GGGCCACAAAATTCGAGGCTCCGAATGCACTGGGCTTGGTGTATTCCTCAAGGAAGGCTGC 420
Db      |||
QY 2112 AAGATTGCGATGTGCCAAGCAGCAGGAGCTGGAGAGGAGGAGGCTGCTCATCTTTGAT 2171
Db      |||
QY 421 AAGATTGCGATGTGCCAAGCAGCAGGAGCTGGAGAGGAGGAGGCTGCTCATCTTTGAT 480
Db      |||
QY 2172 GACTCCTTTGAGCAGAGGATGCGAGATGCGCTCATCTTTCCGGCTGATATTCATCGTG 2231
Db      |||
QY 481 GACTCCTTTGAGCAGAGGATGCGAGATGCGCTCATCTTTCCGGCTGATATTCATCGTG 540
Db      |||
QY 2232 GATGTGTGCTCCCGAACTGACACACAGCAGAGAGCAGGCTTCCAGCAATTTAGCAT 2291
Db      |||
QY 541 GATGTGTGCTCCCGAACTGACACACAGCAGAGAGCAGGCTTCCAGCAATTTAGCAT 600
Db      |||
QY 2292 GAATTCATCAAGCTTGGGAACTCTCG 2319
Db      |||
QY 601 GAATTCATCAAGCTTGGGAACTCTCG 628
Db      |||
```

## RESULT 9

BF336604

LOCUS	BP336604	603 bp	mRNA	linear	EST 17-SEP-2004
DEFINITION	BP336604 Sugano cDNA library, coronary artery smooth muscle cell				
ACCESSION	BP336604				
VERSION	BP336604.1	GI:52266184			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 603)				
JOURNAL	Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,				
COMMENT	Mizushima-Sugano,J., Nakai,K. and Sugano,S.				
FEATURES	Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions Genome Res. 14 (9), 1711-1718 (2004) Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp.				
ORIGIN	Location/Qualifiers				
	1..603				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="SMR06083"				
	/tissue_type="coronary artery"				
	/cell_type="smooth muscle cell"				
	/clone_lib="Sugano cDNA library, coronary artery smooth muscle cell"				
	Query Match 24.4%; Score 568; DB 5; Length 603;				
	Best Local Similarity 100.0%; Pred. No. 2.9e-293;				
	Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	801 GCAGTATATGAACCTCTAGAAAAATGAAGGGATAGAAATCAAGAGTAACTGCTCCCCCT	860			
DB	36 GCAGTATATGAACCTCTAGAAAAATGAAGGGATAGAAATCAAGAGTAACTGCTCCCCCT	95			
QY	861 GAGGATATCCTGTAGAGATTACAGGTAACTCTAGAAGAGTAAGCACTTTTTCCTGTG	920			
DB	96 GAGGATATCCTGTAGAGATTACAGGTAACTCTAGAAGAGTAAGCACTTTTTCCTGTG	155			
QY	921 GAAAGAACAGCAGGAGTAGTACCAACAGAAACAAATGAAAAACAGATGATCAGAACAAAA	980			
DB	156 GAAAGAACAGCAGGAGTAGTACCAACAGAAACAAATGAAAAACAGATGATCAGAACAAAA	215			
QY	981 GAAAAAGTTAGAAAAAGAGCCTAACTTTTAAATAAATTTGATAAGACTATTAAAGCT	1040			
DB	216 GAAAAAGTTAGAAAAAGAGCCTAACTTTTAAATAAATTTGATAAGACTATTAAAGCT	275			
QY	1041 GAACTTGATGTCGAGAAAAACTCCGTAAGGGGAAAAATTTGAGGAGCAGTGAAATGCA	1100			
DB	276 GAACTTGATGTCGAGAAAAACTCCGTAAGGGGAAAAATTTGAGGAGCAGTGAAATGCA	335			
QY	1101 TTTAAAGAACTAGTAGCGAAATACCCTCAGAGTCACAGCAAGATATGGGAAGGCCGAG	1160			
DB	336 TTTAAAGAACTAGTAGCGAAATACCCTCAGAGTCACAGCAAGATATGGGAAGGCCGAG	395			
QY	1161 TGTGAGGATGATTTGGCTCAGAGAGGAGAGTAATGAGTGTCTACGTGGAGCATCGAG	1220			
DB	396 TGTGAGGATGATTTGGCTCAGAGAGGAGAGTAATGAGTGTCTACGTGGAGCATCGAG	455			
QY	1221 ACCTCAAGAGGTGGCCAGCCTACTGATGTCCCTGAGACCTGTCTGAAGCTAGTTTG	1280			
DB	456 ACCTCAAGAGGTGGCCAGCCTACTGATGTCCCTGAGACCTGTCTGAAGCTAGTTTG	515			
QY	1281 AAGCGTCGCTCAGACAGGCAACAAATTTCTAGCTCATATCAGAGGTTTCCCTCTTACCTG	1340			
DB	516 AAGCGTCGCTCAGACAGGCAACAAATTTCTAGCTCATATCAGAGGTTTCCCTCTTACCTG	575			

QY 1522 TTCTGAGAGCATCCCATATTTAAAGGAAGAAATAGATCCGAGATCCTGGCACTGATG 1581  
Db 441 TTCTGAGAGCATCCCATATTTAAAGGAAGAAATAGATCCGAGATCCTGGCACTGATG 500  
QY 1582 ATGGGAGATTTTATTTTCCACCTGGGGATGCCATGACAGAGGTTGGGAACAAAGAGGCAT 1641  
Db 501 ATGGGAGATTTTATTTTCCACCTGGGGATGCCATGACAGAGGTTGGGAACAAAGAGGCAT 560  
QY 1642 ATAAGTGTATGAGCTTGGGCAC 1664  
Db 561 ATAAGTGTATGAGCTTGGGCAC 583

RESULT 11  
AUI10952  
LOCUS  
DEFINITION AUI10952 NT2RP3 Homo sapiens cDNA clone NT2RP3001731 5', mRNA EST 01-AUG-2002  
sequence.  
ACCESSION AUI10952  
VERSION AUI10952  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Ota, I., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,  
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and  
Isogai, T.  
HRI human cDNA project  
Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-1975  
Fax: 81-438-52-1986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- and 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.  
Location/Qualifiers  
1. 859  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP3001731"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/clone\_lib="NT2RP3"  
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"

ORIGIN  
Query Match 23.9%; Score 556; DB 1; Length 859;  
Best Local Similarity 99.7%; Pred. No. 8.8e-287;  
Matches 656; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1013 AAATPAAATTTGATAAGACTTAAAGCTGAACCTTGAATGCTGCGAGAAAACCTCGTAAAG 1072  
Db 1 AAATPAAATTTGATAAGACTTAAAGCTGAACCTTGAATGCTGCGAGAAAACCTCGTAAAG 60  
QY 1073 GGGAAAAATTTGAGAGAGCAGTGAATGCTTAAAGACTAGTACGCAATACCTTCAGAG 1132  
Db 61 GGGAAAGTTGAGAGAGCAGTGAATGCTTAAAGACTAGTACGCAATACCTTCAGAG 120  
QY 1133 TCCACGAGCAGATGGAAGCCGAGTGTGAGGATGATTTGGCTGGAAGAGAGAAG 1192  
Db 121 TCCACGAGCAGATGGAAGCCGAGTGTGAGGATGATTTGGCTGGAAGAGAGAAG 180  
QY 1193 TAATGAGGTGCTAGCTGGAGCCATCGAGACCTACCAAGAGGTGGCCGCTACCTGATGT 1252

Db 181 TAATGAGGTGCTAGCTGGAGCCATCGAGACCTACCAAGAGGTGGCCGCTACCTGATGT 240  
QY 1253 CCTTCAGACCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGG 1312  
Db 241 CCTTCAGACCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGG 300  
QY 1313 TCATATGAGAGGTTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTTCCCAATGATC 1372  
Db 301 TCATATGAGAGGTTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTTCCCAATGATC 360  
QY 1373 TTCTTTAAAAAATGACCTTGGCGTGGGATACCTTTGATAGGAGATATGACATGCANA 1432  
Db 361 TTCTTTAAAAAATGACCTTGGCGTGGGATACCTTTGATAGGAGATATGACATGCANA 420  
QY 1433 GAAAGTTTATGAGAGGTCGCTGAGTGTGACACCTTATGATGCTTCTTAAAGTCCATTA 1492  
Db 421 GAAAGTTTATGAGAGGTCGCTGAGTGTGACACCTTATGATGCTTCTTAAAGTCCATTA 480  
QY 1493 TGGCTTATCTGCTGAAGGCACAGACAAAATTGCTGAGAGCATCCCATATTTAAAGGAAGG 1552  
Db 481 TGGCTTATCTGCTGAAGGCACAGACAAAATTGCTGAGAGCATCCCATATTTAAAGGAAGG 540  
QY 1553 AATGAATCCGAGAGATCCCTGGCACTGATGGGAGATTTTATTTCCACCTGGGGATGC 1612  
Db 541 AATGAATCCGAGAGATCCCTGGCACTGATGGGAGATTTTATTTCCACCTGGGGATGC 600  
QY 1613 CATGCAGAGGTTGGGAACAAGAGGCATATAGTGTGATGCTTGGGCAAGAGA 1670  
Db 601 CATGCAGAGGTTGGGAACAAGAGGCATATAGTGTGATGCTTGGGCAAGAGA 658

RESULT 12  
AW996828/6  
LOCUS  
DEFINITION QV3-BN0047-160300-123-a02 BN0047 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW996828  
VERSION AW996828.1 GI:8257062  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjovschi-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
PUBMED 10737800  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=6t2-QV3-BN0047-160  
300-123-a02&t3=2000-03-16&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 643.  
Location/Qualifiers  
1. 651  
/organism="Homo sapiens"  
/mol\_type="mRNA"

/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BN0047"  
/note="Organ: breast normal; Vector: puc18; Site: 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ORIGIN		Query Match	23.4%; Score 543; DB 2: Length 651;
		Best Local Similarity	100.0%; Pred. No. 8.3e-280;
		Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1366	ATGATCTCTCTTAAATAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACA	1425
DB	620	ATGATCTCTCTTAAATAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACA	561
QY	1426	ATGCAAGAAAGTTTATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAG	1485
DB	560	ATGCAAGAAAGTTTATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAG	501
QY	1486	TCCATTATGGCTTCATCTGAGGCACAGACAAATTTGCTGAGAGCATCCATATTAA	1545
DB	500	TCCATTATGGCTTCATCTGAGGCACAGACAAATTTGCTGAGAGCATCCATATTAA	441
QY	1546	AGGAAGGAATAGATCCGAGATCCTGGCACTGATGGGAGATTTTATTTCCACTGG	1605
DB	440	AGGAAGGAATAGATCCGAGATCCTGGCACTGATGGGAGATTTTATTTCCACTGG	381
QY	1606	GGGATGCCATGACAGAGGTGGGAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACA	1665
DB	380	GGGATGCCATGACAGAGGTGGGAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACA	321
QY	1666	AGAGAGGACACTTTCATCTGTCTGGCAACGCTCACTCTACAATGTGAATGGACTGAAAG	1725
DB	320	AGAGAGGACACTTTCATCTGTCTGGCAACGCTCACTCTACAATGTGAATGGACTGAAAG	261
QY	1726	CACAGCTTTGGTGGACCCCAAGAAACGGGTACACAGTGTAGTAAGTCTTTAGAAA	1785
DB	260	CACAGCTTTGGTGGACCCCAAGAAACGGGTACACAGTGTAGTAAGTCTTTAGAAA	201
QY	1786	GAACTGGAAGTATATCCAGATGAGGCTTGCAGTGTGATGATGAAGCCAAAGGTCTCT	1845
DB	200	GAACTGGAAGTATATCCAGATGAGGCTTGCAGTGTGATGATGAAGCCAAAGGTCTCT	141
QY	1846	TCTGCTGAGGATGAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTACGCTGTGGC	1905
DB	140	TCTGCTGAGGATGAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTACGCTGTGGC	81
QY	1906	AGC 1908	
DB	80	AGC 78	

RESULT 13  
BX403636  
LOCUS  
DEFINITION  
BX403636 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA007ZC04  
5-PRIME, mRNA sequence.  
ACCESSION  
BX403636  
VERSION  
BX403636.2 GI:46924400  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1011)  
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)

COMMENT

On May 13, 2003 this sequence version replaced gi:30635043.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 4537.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CL0BA007ZC04&P=4537.r.

FEATURES  
source

1..1011  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CLOBA007ZC04"  
/issue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo (dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN

Query Match		23.3%;	Score 541;	DB 5;	Length 1011;
Best Local Similarity		99.3%;	Pred. No. 1.1e-278;		
Matches 841;		Conservative	0;	Mismatches 6;	Indels 0; Gaps 0;
QY	1	CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCGGCGGCGGAGAGACAA	120		
DB	110	CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCGGCGGCGGAGAGACAA	169		
QY	61	GCTCCGCGCAGCGGTAGCAGAGTGGCGGAGAGAGCGGCGGCGGCGGCGGAGAGACAA	120		
DB	170	GCTCCGCGCAGCGGTAGCAGAGTGGCGGAGAGAGCGGCGGCGGCGGCGGAGAGACAA	229		
QY	121	AGCATGGAGGACACAAAGATGGGAGAAAGCGGAGCTCTCGGGAACCTTCATTTCTCACGT	180		
DB	230	AGCATGGAGGACACAAAGATGGGAGAAAGCGGAGCTCTCAGGAACCTTCATTTCTCACGT	289		
QY	181	GGTTTATGGTGAATTCGATCTGGCGGTCTGGACATCTGTAGCTGCTGTTGGTTTGTATC	240		
DB	290	GGTTTATGGTGAATTCGATCTGGCGGTCTGGACATCTGTAGCTGCTGTTGGTTTGTATC	349		
QY	241	TTGTTGACTATAGGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG	300		
DB	350	TTGTTGACTATAGGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG	409		
QY	301	ATTTTGTATGGTATGATGCCAAAGTTTATAGGACTTTAAGAGAGATCTACTTCAGAGC	360		
DB	410	ATTTTGTATGGTATGATGCCAAAGTTTATAGGACTTTAAGAGAGATCTACTTCAGAGC	469		
QY	361	CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTCTGTGG	420		
DB	470	CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTCTGTGG	529		
QY	421	AGGCAGAACCCAGAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	480		
DB	530	AGGCAGAACCCAGAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	589		
QY	481	AAATGGTACACGCAACATGTTGAGGAGAGAGCTTGCACAAAGAGATGGACCCACAG	540		
DB	590	AAATGNTACACGCAACATGTTGAGGAGAGAGCTTGCACAAAGAGATGGACCCACAG	649		
QY	541	GAGAACCAACACAAAGAGGATGATGAGTTCTTATGGGAGCTGATGATGATGATGATGATG	600		
DB	650	GAGAACCAACACAAAGAGGATGATGAGTTCTTATGGCGAGCTGATGATGATGATGATGATG	709		
QY	601	AGACCCCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTTACCACGCTGGAGAGA	660		

```

Db      710 AGACCCCTGGAACCTGAAGTATCTTATGAAGAAACCGAGCATAGTTTACCACGTGGAAGAGA 769
Qy      661 CAGTTTTCACAGACTGTAAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720
Db      770 CAGTTTTCACAGACTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 829
Qy      721 ATTCCAGTGAACCACTAGTGAAGATGAAGATGGAAGATGGAAGATGGAAGATGGAAGAT 780
Db      830 ATTCCAGTGAACCACTAGTGAAGATGAAGATGGAAGATGGAAGATGGAAGATGGAAGAT 889
Qy      781 ACCAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATGAATCA 840
Db      890 ACCAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATGAATCA 949
Qy      841 CAGAAAT 847
Db      950 CAGAAAT 956

```

```

RESULT 14
CN258117 589 bp mRNA linear EST 16-MAY-2004
LOCUS 1700531862302 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN258117
ACCESSION CN258117.1 GI:47274531
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 589)
Bradenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandali, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
CONTACT: Bradenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbradenberger@geron.com
Insert Length: 589 Std Error: 0.00.
Location/Qualifiers
1..589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

```

## FEATURES

```

source
1..589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

```

## ORIGIN

```

Query Match 23.1%; Score 538; DB 7; Length 589;
Best Local Similarity 100.0%; Pred. No. 4e-277;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1216 TCAGAGCTTACCAAGAGGTGGCCAGCTTACCTGATGTCCTGAGACCTGCTGAAGCTGA 1275
Db 52 TCAGAGCTTACCAAGAGGTGGCCAGCTTACCTGATGTCCTGAGACCTGCTGAAGCTGA 111
Qy 1276 GTTTGAAGCTTCTCAGACAGCAACAAATTTCTAGTCTATATGAGAGTTCCCTGCTTA 1335
Db 112 GTTTGAAGCTTCTCAGACAGCAACAAATTTCTAGTCTATATGAGAGTTCCCTGCTTA 171
Qy 1336 CCCTGCAGAGATTAGTTCAACTATTTCCTCAATGATATCTTCCATAAATAAGACCTTGGC 1395

```

```

Db      172 CCTCTCAGAGATTAGTTCAACTATTTCCTCAATGATCTTCTTAAAAATGACCTTGGCG 231
Qy      1396 TGGGATACCTCTTGTATAGGAGATAATGACAATGCAAGAAAGTTTATGAAGAGGTGCTGA 1455
Db      232 TGGGATACCTCTTGTATAGGAGATAATGACAATGCAAGAAAGTTTATGAAGAGGTGCTGA 291
Qy      1456 GTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTATGGCTTCATCTCTGAAGGCACAGA 1515
Db      292 GTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTATGGCTTCATCTCTGAAGGCACAGA 351
Qy      1516 ACAAATATGCTGAGAGCATCCCATATTTTAAAGAGGAGTAATAGATCCGGAGATCTCGGCA 1575
Db      352 ACAAATATGCTGAGAGCATCCCATATTTTAAAGAGGAGTAATAGATCCGGAGATCTCGGCA 411
Qy      1576 CTGATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGCGAGAGGTTGGGAAACAAAG 1635
Db      412 CTGATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGCGAGAGGTTGGGAAACAAAG 471
Qy      1636 AGGCATATAAGTGTATGAGCTTGGGCAACAAGAGAGGACACTTTGCACTCTGTCTGGCAAC 1695
Db      472 AGGCATATAAGTGTATGAGCTTGGGCAACAAGAGAGGACACTTTGCACTCTGTCTGGCAAC 531
Qy      1696 GCTCACTCTCAATGTGAATGGACTGAAAGCAGACGCTTGTGTGACCCCAAAAGAAAC 1753
Db      532 GCTCACTCTCAATGTGAATGGACTGAAAGCAGACGCTTGTGTGACCCCAAAAGAAAC 589

```

## RESULT 15

```

BP266690 581 bp mRNA linear EST 16-SEP-2004
LOCUS BP266690 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
DEFINITION JTH04779, mRNA sequence.
ACCESSION BP266690
VERSION BP266690.1 GI:52181921
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 581)
Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
CONTACT: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH04779"
/tissue_type="thyroid"
/cell_line="JTH"
/note="Thyroid tumor"

```

## FEATURES

```

source
1..581

```

## ORIGIN

```

Query Match 23.1%; Score 537; DB 5; Length 581;
Best Local Similarity 100.0%; Pred. No. 1.4e-276;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1158 CAGTGTGAGATGATTTGGCTGAGAAGAGAGAGTAATGAGTCTGCTGAGGAGCATC 1217
Db 45 CAGTGTGAGATGATTTGGCTGAGAAGAGAGAGTAATGAGTCTGCTGAGGAGCATC 104
Qy 1218 GAGACCTACCAAGAGTGGCCAGCTTACCTGATGTCCTTGCAGACCTCTCTGAAGCTGAT 1277

```



Db	105	GAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCCTGCAGACCTGCTGAAGCTGAGT	164
Qy	1278	TTGAAGCGTGGCTCAGACAGGCAACAATTTCTAGGTCAATGAGAGGTTCCCTGCTTACC	1337
Db	165	TTGAAGCGTGGCTCAGACAGGCAACAATTTCTAGGTCAATGAGAGGTTCCCTGCTTACC	224
Qy	1338	CTGCAGAGATAGTTCAACTATTTCCTCAATGATCTTCCCTTAAATAATGACCTTGGCGTG	1397
Db	225	CTGCAGAGATAGTTCAACTATTTCCTCAATGATCTTCCCTTAAATAATGACCTTGGCGTG	284
Qy	1398	GGATACCTCTTGATAGAGATATGACAAATGCAAGAAAGTTTATGAAGAGGTGCTGAGT	1457
Db	285	GGATACCTCTTGATAGAGATATGACAAATGCAAGAAAGTTTATGAAGAGGTGCTGAGT	344
Qy	1458	GTGCACACTAATGATGGCTTGTCTTAAAGTCAATGATGCTTCCCTGAAGGACACAGAAC	1517
Db	345	GTGCACACTAATGATGGCTTGTCTTAAAGTCAATGATGCTTCCCTGAAGGACACAGAAC	404
Qy	1518	AAATTTGCTGAGACATCCCATATTTTAAAGGAAGGAATAGAAATCCGGAGATCTCTGGCACT	1577
Db	405	AAATTTGCTGAGACATCCCATATTTTAAAGGAAGGAATAGAAATCCGGAGATCTCTGGCACT	464
Qy	1578	GATGATGGAGATTTTATTTCCACTCTGGGGATGCCATGACAGAGGGTTGGGAACAAAGAG	1637
Db	465	GATGATGGAGATTTTATTTCCACTCTGGGGATGCCATGACAGAGGGTTGGGAACAAAGAG	524
Qy	1638	GCATATAAGTGGTATGAGCTTGGGCAAGAGAGGACACTTTTGCATCTGTCTGGCAA	1694
Db	525	GCATATAAGTGGTATGAGCTTGGGCAAGAGAGGACACTTTTGCATCTGTCTGGCAA	581
RESULT 16			
BP270196			
LOCUS	BP270196	582 bp	mRNA linear EST 16-SEP-2004
DEFINITION	clone KAR01529, mRNA sequence.		
ACCESSION	BP270196		
VERSION	BP270196.1	GI:52185428	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Suzuki,Y., Yanashita,R., Shiota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S. Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions Genome Res. 14 (9), 1711-1718 (2004)		
AUTHORS	Contact: Yutaka Suzuki		
TITLE	Department of Virology		
JOURNAL	Institute of Medical Science, University of Tokyo		
COMMENT	4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan		
FEATURES	Email: yusuzuki@ims.u-tokyo.ac.jp.		
Source	1. .582		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="KAR01529"		
	/tissue_type="small intestine"		
	/clone_lib="Sugano cdna library, small intestine"		
ORIGIN			
Query Match	22.8%; Score 531, DB 5; Length 582;		
Best Local Similarity	99.8%; Pred. No. 2.3e-273;		
Matches 581; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Qy	1651	ATGAGCTTGGGCACAGAGAGGACACTTTTGCATCTGCTGGCAACGCTCACTCAATG	1710
Db	1	ATGAGCTTGGGCACAGAGAGGACACTTTTGCATCTGCTGGCAACGCTCACTCAATG	60
Qy	1711	TGAATGACTGAAAGCAGCAGCTTGGTGGACCCCAAGAAAGAGGGCTACAGAGTTAG	1770
Db	61	TGAATGACTGAAAGCAGCAGCTTGGTGGACCCCAAGAAAGAGGGCTACAGAGTTAG	120
Qy	1771	TAAAGCTTTTGAAGAAACCTGGAAGTTAATCCGAGATGAAGGCCTTCAGTGTGAGATA	1830
Db	121	TAAAGCTTTTGAAGAAACCTGGAAGTTAATCCGAGATGAAGGCCTTCAGTGTGAGATA	180
Qy	1831	AAGCCAAAGGCTCTTTCCTGCTGAGGATGAAAACCTCAGGGAAAAAGGGGACTGGAGCC	1890
Db	181	AAGCCAAAGGCTCTTTCCTGCTGAGGATGAAAACCTCAGGGAAAAAGGGGACTGGAGCC	240
Qy	1891	AGTTACGCTGTGGCAGCAGGAAGAAGAAATGAAAATGCCCTGCAAGAGAGCTCCTAAAA	1950
Db	241	AGTTACGCTGTGGCAGCAGGAAGAAGAAATGAAAATGCCCTGCAAGAGAGCTCCTAAAA	300
Qy	1951	CTGTGATCTTACTAGAAAAGTTTCCCGAGACAAACAGGATGCAGAAAGAGGACAGATCAAT	2010
Db	301	CTGTGATCTTACTAGAAAAGTTTCCCGAGACAAACAGGATGCAGAAAGAGGACAGATCAAT	360
Qy	2011	ATTCCATCATGACACCCCGGACTCAGCTGTGGCCGACACAGGGCCCAAACTGCAGGC	2070
Db	361	ATTCCATCATGACACCCCGGACTCAGCTGTGGCCGACACAGGGCCCAAACTGCAGGC	420
Qy	2071	TCCGAATGCACCTGGGCTTGGTATTCCCAAGAGAGGCTGCAAGATTCGATGTGCCAAG	2130
Db	421	TCCGAATGCACCTGGGCTTGGTATTCCCAAGAGAGGCTGCAAGATTCGATGTGCCAAG	480
Qy	2131	AGACCAAGACCTGGGAGAGGCAAGGTGCTCATCTTTGATGACTCTTTGATGACTCTTTGAGCAGG	2190
Db	481	AGACCAAGACCTGGGAGAGGCAAGGTGCTCATCTTTGATGACTCTTTGATGACTCTTTGAGCAGG	540
Qy	2191	TATGGCAGGATGCTCATCTTTCCGGCTGATTTATCTGTTG 2232	
Db	541	TATGGCAGGATGCTCATCTTTCCGGCTGATTTATCTGTTG 582	
RESULT 17			
CD643948			
LOCUS	CD643948	765 bp	mRNA linear EST 17-JUN-2003
DEFINITION	AGENCOURT 14550808 NIA Human H1 Embryonic Stem Cell cdna library (Long) Homo sapiens cdna clone IMAGE:30425975 5', mRNA sequence.		
ACCESSION	CD643948		
VERSION	CD643948.1	GI:31815344	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 765)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Daniela S. Gerhard, Ph.D.		
	Office of Cancer Genomics		
	National Cancer Institute / NIH		
	Bldg. 31 Rm10A07 Bethesda, MD 20892		
	Email: csapsb-remail.nih.gov		
	Tissue Procurement: Irene Ginis and Mahendra Rao, NIA		
	cdna Library Preparation: Yulan Piao and Minoru KO		
	cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC c		
	can be found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: NDAM511 row: a column: 24		
	High quality sequence stop: 605.		
FEATURES	Location/Qualifiers		
source	1. .765		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:30425975"		
	/tissue_type="Embryonic Stem cells"		

/cell\_line="WA01"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA Library  
(long)"  
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;  
This is a long-transcript enriched cDNA library (Genome  
Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01/H1  
cell line. Undifferentiated human ES cell line WA01/H1  
was obtained from WiCell Research Institute, Inc.,  
Madison, WI, cultured according to their instructions, on  
MEF feeders. They formed round colonies with defined edges  
and were positive for alkaline phosphatase, SSEA-4, OCT3,  
OCT4, REK1, UTF, TERT, SOX2, CX43 and CX45. They are  
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,  
TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days  
after plating), the ES cells from 4 X 6cm dishes were  
treated with 1 mg/ml collagenase, type IV  
(Invitrogen/GIBCO) for 5-10 min and gently scraped off  
with 5 ml pipette. RNA was purified with TRIzol Reagent  
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558  
(2001). [PMID: 11544199]) Double-stranded cDNAs were  
synthesized with an Oligo(dT) primer [Invitrogen].  
5'-PGACTAGTCTAGATCGAGCGGCGCCCTTTTCTTTTCTTTT-3' from  
3.4% of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to lona-linker Lr-Sal4, purified by phenol/chloroform  
extraction, and separated from free linkers by  
Centricon-100 column. Then, the cDNAs were amplified by  
long-range high fidelity PCR using Ex Taq polymerase  
(Takara) with a primer Sal4-S for 25 cycles. The products  
were purified by phenol/chloroform extraction and  
Centricon-100 column. The cDNAs were digested with SalI  
and NotI enzymes and cloned into SalI/NotI site of  
pCMV-SPORT6 plasmid vector. The average insert size is  
about 3.6kb."

## ORIGIN

Query Match 22.8%; Score 529; DB 6; Length 765;  
Best Local Similarity 99.8%; Pred. No. 2.9e-272;  
Matches 579; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 404 GGAGCAGGTTCTCTGTGAGCGAGAACCCGACATATCGAAGTGAAGCAAAAGACAAT 463  
DB 12 GGAGCAGGTTCTCTGTGAGCGAGAACCCGACATATCGAAGTGAAGCAAAAGACAAT 71

QY 464 TCAGTCCCTTCTCCATGAATGGTACACGACAGACATGTTGAGGGAGAGACTTGAACA 523  
DB 72 TCAGTCCCTTCTCCATGAATGGTACACGACAGACATGTTGAGGGAGAGACTTGAACA 131

QY 524 AGAAGATGGACCCACAGGAGAACCAACACAGAGGATGATGAGTTCCTTATGGCAGTGA 583  
DB 132 AGAAGATGGACCCACAGGAGAACCAACACAGAGGATGATGAGTTCCTTATGGCAGTGA 191

QY 584 TGTAGATGATGATTTGACACCTCGAACCTGAGTATCTCATGAAGAAACCGACATAG 643  
DB 192 TGTAGATGATGATTTGACACCTCGAACCTGAGTATCTCATGAAGAAACCGACATAG 251

QY 644 TTACCACTGGAGAGACAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTC 703  
DB 252 TTACCACTGGAGAGACAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTC 311

QY 704 TGAGCAGGAAAAATCCAGATTCCAGTGAACAGTGTAGTAGAGATGAAGATTGCACATGA 763  
DB 312 TGAGCAGGAAAAATCCAGATTCCAGTGAACAGTGTAGTAGAGATGAAGATTGCACATGA 371

QY 764 TACAGATGATGTAAACATCCAGTCTATGAGAACAGAGTATATGAACCTCTAGAAAA 823  
DB 372 TACAGATGATGTAAACATCCAGTCTATGAGAACAGAGTATATGAACCTCTAGAAAA 431

QY 824 TGAAGGGATAGAAATCAGAGAGTAACTGTCTCCCTCGAGGATTAATCTCTGAGAGATTC 883  
DB 432 TGAAGGGATAGAAATCAGAGAGTAACTGTCTCCCTCGAGGATTAATCTCTGAGAGATTC 491

QY 884 ACAGGTAATTCAGAGAAAGTAAGCATTTTCTCTGTGGAAGACAGCAGGAAGTACCACC 943  
DB 492 ACAGGTAATTCAGAGAAAGTAAGCATTTTCTCTGTGGAAGACAGCAGGAAGTACCACC 551

QY 944 AGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCA 983  
DB 552 AGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCA 591

RESULT 18  
BX479249 517 bp mRNA linear EST 04-SEP-2003  
LOCUS DKFZP686L22209.r1.686 (synonym: hlc3) Homo sapiens cDNA clone  
DEFINITION DKFZP686L22209.5, mRNA sequence.  
ACCESSION BX479249  
VERSION BX479249.1 GI:31914859  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (Bases 1 to 517)  
AUTHORS Bloeker,H., Boecker,M., Mewes,H.W., Weil,B., Amid,C., Oeanger,A.,  
Fobo,G., Han,M. and Wiemann,S.  
TITLE EST (Bloeker,H., Boecker,M., Mewes,H.W., Weil,B., Amid,C., et al.)  
JOURNAL Unpublished (2003)  
COMMENT Contact: MIPS  
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Sponsored by GFR (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.  
This clone (DKFZP686L22209) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..517  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZP686L22209"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiIb;  
cDNA-collection"

ORIGIN

Query Match 22.2%; Score 517; DB 5; Length 517;  
Best Local Similarity 100.0%; Pred. No. 7.8e-266;  
Matches 517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 926 ACAGCAGGAGTACACACAGAACAAATAGAAAACAGATGATCCAGAACAAAGCAA 985  
DB 1 ACAGCAGGAGTACACACAGAACAAATAGAAAACAGATGATCCAGAACAAAGCAA 60

QY 986 AGTTAGAAAAGAGAGCCCTTAACTTTTAAATAAATTTGATAAGACTATTAAGCTGA 1045  
DB 61 AGTTAGAAAAGAGAGCCCTTAACTTTTAAATAAATTTGATAAGACTATTAAGCTGA 120

QY 1046 TGATGCTCAGAAAAAATCCCGTAAAGGGGAAAAATTTGAGGAAGCAGTGAATGCAATTAA 1105  
DB 121 TGATGCTCAGAAAAAATCCCGTAAAGGGGAAAAATTTGAGGAAGCAGTGAATGCAATTAA 180

QY 1106 AGAATCTAGTACCAAAATACCTCAGAGTCCAGCAGCAAGATATGGAGGCGCAGTGA 1165  
DB 181 AGAATCTAGTACCAAAATACCTCAGAGTCCAGCAGCAAGATATGGAGGCGCAGTGA 240

QY 1166 GGATGATTTGGCTGGAAGAGGAGAGTAATGAGGTGTACGTGAGGCCATCGAGACCTA 1225

Db	241	GGATGATTTGGCTGAGAAGAGGAGAAATGATGAGTCTACGTGGAGCCATCGAGACCTA	300
Qy	1226	CCAAGAGGTGGCCAGCTTACCTGATGTCCTCGCAGACCTGCTGAAGCTGAGTTTGAAGCG	1285
Db	301	CCAAGAGGTGGCCAGCTTACCTGATGTCCTCGCAGACCTGCTGAAGCTGAGTTTGAAGCG	360
Qy	1286	TCGCTCAGACGAGCAACAATTTCTAGGTCTATGATGAGAGGTTCCCTGCTTACCCCTGCAGAG	1345
Db	361	TCGCTCAGACGAGCAACAATTTCTAGGTCTATGATGAGAGGTTCCCTGCTTACCCCTGCAGAG	420
Qy	1346	ATTAGTTCAACTATTTCCCAATGATCTTCTTAAAAAATGACTTGGCGCTGGGATACCT	1405
Db	421	ATTAGTTCAACTATTTCCCAATGATCTTCTTAAAAAATGACTTGGCGCTGGGATACCT	480
Qy	1406	CTTGATAGGAGATATGACAAATGCAAAAGAAAGTTTAT	1442
Db	481	CTTGATAGGAGATATGCAATGCAAAAGAAAGTTTAT	517
RESULT 19			
BI333062			
LOCUS	602996326F1 NIH_MGC_12	721 bp mRNA linear	EST 30-JUL-2001
DEFINITION	602996326F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5138622 5', mRNA sequence.		
ACCESSION	BI333062		
VERSION	BI333062.1	GI:15017719	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)		
TITLE	Unpublished (1999)		
JOURNAL	Contact: Robert Strausberg, Ph.D.		
COMMENT	Email: cgaabs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LL4W1341 row: c column: 07 High quality sequence stop: 719. Location/Qualifiers		
FEATURES	1..721		
source	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5138622"		
	/tissue_type="cervical carcinoma cell line"		
	/lab_host="DH10B"		
	/clone_lib="NIH MGC 12"		
	/note="Organ: Cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 Kb. Library prepared by Life Technologies."		
ORIGIN	1..721		
Query Match	21.8%;	Score 507;	DB 4; Length 721;
Best Local Similarity	99.8%;	Pred. No. 2e-260;	
Matches 557;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	CGGACCGTGCATATGGCCCGAGCGTGAAGTCCCAAGAGAGCGCGCAACACAGCAGCGG	60
Db	113	CGGACCGTGCATATGGCCCGAGCGTGAAGTCCCAAGAGAGCGCGCAACACAGCAGCGG	172
Qy	61	GCTCCGGCAGCGGTAGCAGGTCGGGCGAGCAGAGCCCGGGCGCCCGGAGAGACAA	120
Db	173	GCTCCGGCAGCGGTAGCAGGTCGGGCGAGCAGAGCCCGGGCGCCCGGAGAGACAA	232

Qy	121	AGCATGGAGGACACAAAGAAATGGAGGAAAGCGGACTCTCGGAACTTTCATTTCTTCAGT	180
Db	233	AGCATGGAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCATTTCTTCAGT	292
Qy	181	GGTTTATGGTATGATTCATTTGCTGGCGCTGCGACATCTGTAGCTGCTGTTTGGTTTGATC	240
Db	293	GGTTTATGGTATGATTCATTTGCTGGCGCTGCGACATCTGTAGCTGCTGTTTGGTTTGATC	352
Qy	241	TTGTTGATCTATGAGGAAGTTCTAGGAAAACCTAGGAATCTATGATCTGATGGTATGAGG	300
Db	353	TTGTTGATCTATGAGGAAGTTCTAGGAAAACCTAGGAATCTATGATCTGATGGTATGAGG	412
Qy	301	ATTTTATGATGATGATGATGCAAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360
Db	413	ATTTTATGATGATGATGATGCAAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC	472
Qy	361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGTTCTGTGG	420
Db	473	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGTTCTGTGG	532
Qy	421	AGCGAAGCCCGCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	480
Db	533	AGCGAAGCCCGCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	592
Qy	481	AAATGTTACACCGCAGAACATGTTGAGGGAGAGACTTGGCAACAAGAGATGACCCACAG	540
Db	593	AAATGTTACACCGCAGAACATGTTGAGGGAGAGACTTGGCAACAAGAGATGACCCACAG	652
Qy	541	GAGAACCAACACAGAGG 558	
Db	653	GAGAACCAACACAGAGG 670	
RESULT 20			
BI333062			
LOCUS	17000418219954 GRN_EB	711 bp mRNA linear	EST 16-MAY-2004
DEFINITION	17000418219954 GRN_EB Homo sapiens cDNA 5', mRNA sequence.		
ACCESSION	BI333062		
VERSION	BI333062.1	GI:47368766	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, Y., Xu, C., Fang, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Brandenberger, R., Guegler, K., Rao, M.S., Mandalam, R., Lebrowski, J and Stanton, L.W.		
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation		
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)		
COMMENT	Contact: Brandenberger R Regenerative Medicine Genon Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@genon.com Insert Length: 711 Std Error: 0.00.		
FEATURES	1..711		
source	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"		
	/clone_lib="GRN_EB"		
	/note="Oligo dt primed, full-length enriched cDNA library from embryoid body outgrowths derived from HES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."		
ORIGIN			

Query Match 21.2%; Score 493; DB 7; Length 711;  
Best Local Similarity 99.8%; Pred. No. 6.8e-253;  
Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACCGTCCATGCGCCAGCGTAAGAAATCCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 168 CGGACCGTCCATGCGCCAGCGTAAGAAATCCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 227  
QY 61 GCTCCGGCAGCGGTAGCAGCAGTCCGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
DB 228 GCTCCGGCAGCGGTAGCAGCAGTCCGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 287  
QY 121 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCATCTTCACGT 180  
DB 288 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCATCTTCACGT 347  
QY 181 GGTATGCTGATGTCATTTGCTGGGCGTCTGGACATCTGTAGCTGTGCTTTGTTGATC 240  
DB 348 GGTATGCTGATGTCATTTGCTGGGCGTCTGGACATCTGTAGCTGTGCTTTGTTGATC 407  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATCGAG 300  
DB 408 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATCGAG 467  
QY 301 ATTGTTGATGCTGATGTCATTTGCTGGGCGTCTGGACATCTGTAGCTGTGCTTTGTTGATC 360  
DB 468 ATTGTTGATGCTGATGTCATTTGCTGGGCGTCTGGACATCTGTAGCTGTGCTTTGTTGATC 527  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG 420  
DB 528 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG 587  
QY 421 AGGCAAGACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCTCATG 480  
DB 588 AGGCAAGACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCTCATG 647  
QY 481 AAATGTTACGCGACACATGTTGAGGAGAGACTTCCACAGAGAGATGGACCCACAG 540  
DB 648 AAATGTTACGCGACACATGTTGAGGAGAGACTTCCACAGAGAGATGGACCCACAG 707  
QY 541 GAGA 544  
DB 708 GAGA 711

## RESULT 21

BQ232651  
LOCUS  
DEFINITION  
AGENCOURT 7574911 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6059910  
5', mRNA sequence.

ACCESSION  
BQ232651

VERSION  
BQ232651.1 GI:20414051

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/.

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: DCTP/DTP/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM13327 row: f column: 07

High quality sequence stop: 667.

## FEATURES

source

Location/Qualifiers

1..868

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6059910"

/tissue\_type="large cell carcinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 68"

/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.8 kb. Library constructed by Life

Technologies."

## ORIGIN

Query Match

Best Local Similarity

Matches 491; Conservative

21.1%; Score 491; DB 5; Length 868;

100.0%; Pred. No. 8.4e-252;

Mismatches 0; Indels 0; Gaps 0;

QY 456 GAACAATTCAGTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGAGAGAC 515

DB 11 GAACAATTCAGTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGAGAGAC 70

QY 516 TTGCAACAAGAAGATGGACCCACAGAGAACCCACAAAGAGGATGATGATTTCTTATG 575

DB 71 TTGCAACAAGAAGATGGACCCACAGAGAACCCACAAAGAGGATGATGATTTCTTATG 130

QY 576 GCGACTGATGTAGATGATGATTTGAGACCTTGGAACTGAACTGATCTCATGAGAAC 635

DB 131 GCGACTGATGTAGATGATGATTTGAGACCTTGGAACTGAACTGATCTCATGAGAAC 190

QY 636 GACCATGATTTACACCTGGAAGAGACAGTTTCAAGACTGTATCAGGATATGGAAG 695

DB 191 GACCATGATTTACACCTGGAAGAGACAGTTTCAAGACTGTATCAGGATATGGAAG 250

QY 696 ATGATGTCTGAGCAGGAAATCCAGATTCAGTGAACCCAGTAGTAGAAGATGAAAGATTG 755

DB 251 ATGATGTCTGAGCAGGAAATCCAGATTCAGTGAACCCAGTAGTAGAAGATGAAAGATTG 310

QY 756 CACCATGATGATGATGATTAACATACCAAGTCTATCAGGAAACAGCAGTATGAACT 815

DB 311 CACCATGATGATGATGATTAACATACCAAGTCTATCAGGAAACAGCAGTATGAACT 370

QY 816 CTAGAAAATGAAGGGATGAAATCACAGAGTAACTCTCCCTCAGGATATCTCTGTA 875

DB 371 CTAGAAAATGAAGGGATGAAATCACAGAGTAACTCTCCCTCAGGATATCTCTGTA 430

QY 876 GAGATTCACAGTAAATTTGTAGAGAAAGTAAAGCAATTTTCTCTGTGGAAGAACAGCAGGAA 935

DB 431 GAGATTCACAGTAAATTTGTAGAGAAAGTAAAGCAATTTTCTCTGTGGAAGAACAGCAGGAA 490

QY 936 GTACCAACCAGA 946

DB 491 GTACCAACCAGA 501

RESULT 22

CN368829

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 520)

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,

Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,

Lebkowski, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

## TITLE

JOURNAL COMMENT  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
Contact: Brandenberger R  
Regenerative Medicine  
Geront Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 520 Std Error: 0.00.

FEATURES  
source  
1. 520  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, DMSO-treated H9 cell line"  
/clone\_lib="GRN PREHEP"  
/note="oligo dt primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN  
Query Match 20.5%; Score 476; DB 7; Length 520;  
Best Local Similarity 100.0%; Pred. No. 9.1e-244;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 431 CCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATGAAATGGTACA 490  
Db 7 CCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATGAAATGGTACA 66  
QY 491 CGCAGAACATGTTGAGGAGAGAGACTTGCACAAAGAGATGGACCCACAGAGAACACACA 550  
Db 67 CGCAGAACATGTTGAGGAGAGAGACTTGCACACAGAGATGGACCCACAGAGAACACACA 126  
QY 551 ACAAGAGGATGATGATGTTCTTATGCGCACTGATGATGATGATGATGATGATGATGATG 610  
Db 127 ACAAGAGGATGATGATGTTCTTATGCGCACTGATGATGATGATGATGATGATGATGATG 186  
QY 611 ACCTGAAGTATCTGATGAGAACCCGAGCATAGTTTACCAGTGGAGAGACAGATTTTACA 670  
Db 187 ACCTGAAGTATCTGATGAGAACCCGAGCATAGTTTACCAGTGGAGAGACAGATTTTACA 246  
QY 671 AGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAGAAATCCAGATTTCCAGTGA 730  
Db 247 AGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAGAAATCCAGATTTCCAGTGA 306  
QY 731 ACCAGTAGTAGAAGATGAAGATTTGACCATGATACAGATGATGATGATGATGATGATGATGAT 790  
Db 307 ACCAGTAGTAGAAGATGAAGATTTGACCATGATACAGATGATGATGATGATGATGATGATGAT 366  
QY 791 TGAGGAACAGCAGATATGAACTCTAGAAATGAAGGATAGAAATCAGAGAGTAAC 850  
Db 367 TGAGGAACAGCAGATATGAACTCTAGAAATGAAGGATAGAAATCAGAGAGTAAC 426  
QY 851 TGCTCCCTCCCTGAGGATTAATCTGTTAGAGATTCACAGGTAATGTTAGAGAGTA 906  
Db 427 TGCTCCCTCCCTGAGGATTAATCTGTTAGAGATTCACAGGTAATGTTAGAGAGTA 482

RESULT 23  
CN368823  
LOCUS 770 bp mRNA linear EST 16-MAY-2004  
DEFINITION 328775548 GRN\_ES Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION CN368823  
VERSION CN368823.1 GI:47368757  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 770)  
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,

TITLE  
JOURNAL  
COMMENT  
Lebkowski, J and Stanton, L.W.  
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
Contact: Brandenberger R  
Regenerative Medicine  
Geront Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 770 Std Error: 0.00.

FEATURES  
source  
1. 770  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, cell lines H1, H7, and H9"  
/clone\_lib="GRN ES"  
/note="oligo dt primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN  
Query Match 19.8%; Score 461; DB 7; Length 770;  
Best Local Similarity 99.8%; Pred. No. 1.2e-235;  
Matches 511; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGCACCGTCCCAATGCCAGCGTAAGCAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60  
Db 186 CGCACCGTCCCAATGCCAGCGTAAGCAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 245  
QY 61 GCTCCCGCAGCGTAGCAGAGTGCAGGAGCAGAGCCCGGGCCCGGAGAGACAA 120  
Db 246 GCTCCCGCAGCGTAGCAGAGTGCAGGAGCAGAGCCCGGGCCCGGAGAGACAA 305  
QY 121 AGCATGAGAGACACAAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
Db 306 AGCATGAGAGACACAAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 365  
QY 181 GGTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
Db 366 GGTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 425  
QY 241 TTGTTGATGAGAGAGAGTCTAGGAG 300  
Db 426 TTGTTGATGAGAGAGAGTCTAGGAG 485  
QY 301 ATTTTGTGATGAGAGAGAGTCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
Db 486 ATTTTGTGATGAGAGAGAGTCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 545  
QY 361 CAGCAGTCCCGCCAG 420  
Db 546 CAGCAGTCCCGCCAG 605  
QY 421 AGGAG 480  
Db 606 AGGAG 665  
QY 481 AAATGGTACAGCAG 512  
Db 666 AAATGGTACAGCAG 697

RESULT 24  
BG720220  
LOCUS 689 bp mRNA linear EST 08-MAY-2001  
DEFINITION 602692270F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4824712 5', mRNA sequence.  
ACCESSION BG720220  
VERSION BG720220.1 GI:13999407

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing Arrayed by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10736 Row: k Column: 17
High quality sequence stop: 685.
FEATURES
source
Location/Qualifiers
1..689
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4824712"
/lab_host="DH10B"
/clone_lib="NIH MGC 97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: Sali-XhoI
(gtccag); Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to R0F 5; this is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGM/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."
ORIGIN
Query Match 19.7%; Score 458; DB 4; Length 689;
Best Local Similarity 99.8%; Pred. No. 4.6e-234;
Matches 508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGCAGCAGCGCCACAGCAGCAGCG 60
DB 141 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGCAGCAGCGCCACAGCAGCAGCG 200
QY 61 GCTCCGCGCAGCGTAGCAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB GCTCCGCGCAGCGTAGCAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 260
QY 121 AGCATGAGGACACAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 261 AGCATGAGGACACAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 320
QY 181 GGTATTATGTTGATTCCTGGCGCTCTGGACATCTGTAGCTGTCTGTCTGTCTGT 240
DB 321 GGTATTATGTTGATTCCTGGCGCTCTGGACATCTGTAGCTGTCTGTCTGTCTGT 380
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGAG 300
DB 381 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGAG 440
QY 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
DB 441 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 500
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGAGGAG 420
DB 501 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGAGGAG 560

```

---

```

QY 421 AGGCGAAGACCCAGCAATATCGAAGATGAAGCAAGAAAGAAACAAATTCAGTCCCTCTCCATG 480
DB 561 AGGCGAAGACCCAGCAATATCGAAGATGAAGCAAGAAAGAAACAAATTCAGTCCCTCTCCATG 620
QY 481 AAATGGTACACGCGAGAACATGTTGAGGGA 509
DB 621 AAATGGTACACGCGAGAACATGTTGAGGGA 649
RESULT 25
CN274513 455 bp mRNA linear EST 16-MAY-2004
LOCUS 1700600026439 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN274513
ACCESSION CN274513
VERSION CN274513.1 GI:47290927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 455)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R.,
Lebkowski, J. and Stanton, L. W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 455 Std Error: 0.00.
Location/Qualifiers
1..455
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell
line"
/clone_lib="GRN_PHEHEP"
/note="oligo dt primed, full-length enriched cDNA library
from DMSO-treated HES cell line H9 (p22) maintained in
feeder-free conditions"
FEATURES
source
Query Match 19.6%; Score 455; DB 7; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.8e-232;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1059 AAATCCGTAAGAGGGGAAAATTTGAGAGAGCAGTGAATGCAATTTAAAGAACTAGTAGCG 1118
DB 1 AAATCCGTAAGAGGGGAAAATTTGAGAGAGCAGTGAATGCAATTTAAAGAACTAGTAGCG 60
QY 1119 AAATACCTCCAGAGTCCACGAGCAAGATATGGAGAGCGGAGTGTGAGATGATTTGGCT 1178
DB 61 AAATACCTCCAGAGTCCACGAGCAAGATATGGAGAGCGGAGTGTGAGATGATTTGGCT 120
QY 1179 GAGAAGAGAGAGTAATGAGTGCTACGTGGAGCCATCGAGACCTCAAGAGAGTGCGCC 1238
DB 121 GAGAAGAGAGAGTAATGAGTGCTACGTGGAGCCATCGAGACCTCAAGAGAGTGCGCC 180
QY 1239 AGCTTACCTGATGTCCTTCGAGACCTCTCTGAAGCTGAGTTTGAAGCGTCTGAGACAGG 1298
DB 181 AGCTTACCTGATGTCCTTCGAGACCTCTCTGAAGCTGAGTTTGAAGCGTCTGAGACAGG 240
QY 1299 CAACAATTTCTAGGTTCATATGAGAGGTTCCCTGCTTACCTTCAGAGATTAGTTCACTA 1358
DB 241 CAACAATTTCTAGGTTCATATGAGAGGTTCCCTGCTTACCTTCAGAGATTAGTTCACTA 300
QY 1359 TTTCCTCAATGATACCTTCCTTAAATGACCTTTGGCGTGGGATACCTCTTATAGGAGAT 1418

```

Db 301 TTTCCCAATGATATCTTCCTTAAATAATGACCTTGGGTGGATACCTCTTGATAGAGAT 360

Qy 1419 AATGACAATGCAAGAAGTTTATGAGAGGTCTGAGTGTGACACCTAATGATGCTTT 1478

Db 361 AATGACAATGCAAGAAGTTTATGAGAGGTCTGAGTGTGACACCTAATGATGCTTT 420

Qy 1479 GCTAAGTCCATTATGCTTCATCCTGAAGGCACA 1513

Db 421 GCTAAGTCCATTATGCTTCATCCTGAAGGCACA 455

RESULT 26

BP328261 583 bp mRNA linear EST 17-SEP-2004

LOCUS BP328261 Sugano cDNA library, rectum Homo sapiens cDNA clone

DEFINITION RCT00312, mRNA sequence.

ACCESSION BP328261.1 GI:52257341

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 583)

AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES

source 1..583

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="RCT00312"

/tissue\_type="rectum"

/clone\_lib="Sugano cDNA library, rectum"

ORIGIN

Query Match 19.4%; Score 452; DB 5; Length 583;

Best Local Similarity 99.6%; Pred. No. 7.6e-231; Indels 0; Gaps 0;

Matches 552; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1366 ATGATACTTCCCTTAAATAATGACCTTGGGTGGATACCTCTTGATAGAGATATGACA 1425

Db 15 ATGATACTTCCCTTAAATAATGACCTTGGGTGGATACCTCTTGATAGAGATATGACA 74

Qy 1426 ATGCAAGAAGTTTATGAAGGTGCTGAGTGTGACACCTAATGATGCTTGTGTAAG 1485

Db 75 ATGCAAGAAGTTTATGAAGGTGCTGAGTGTGACACCTAATGATGCTTGTGTAAG 134

Qy 1486 TCCATTATGCTTCATCCTGAGGACACAAATTCCTGAGAGCATCCCATATTTAA 1545

Db 135 TCCATTATGCTTCATCCTGAGGACACAAATTCCTGAGAGCATCCCATATTTAA 194

Qy 1546 AGGAAGGATAGAAATCCGGAGATCCTGGCACTGATGGGAGATTTATTTCCACTGG 1605

Db 195 AGGAAGGATAGAAATCCGGAGATCCTGGCACTGATGGGAGATTTATTTCCACTGG 254

Qy 1606 GGATGCCATGACAGAGGTGGCAAAAGAGGCATATAGTGGTATGAGCTTGGGCACA 1665

Db 255 GGATGCCATGACAGAGGTGGCAAAAGAGGCATATAGTGGTATGAGCTTGGGCACA 314

Qy 1666 AGAGAGGACACTTTGATCTGTCTGGCAACGCTCACTCTACATGTGAATGGACTGAAG 1725

Db 315 AGAGAGGACACTTTGATCTGTCTGGCAACGCTCACTCTACATGTGAATGGACTGAAG 374

Qy 1726 CACAGCCTTGGTGGACCCCAAAAGAAACGGGCTACAGAGTTAGTAAGTCTTTAGAAA 1785

Db 375 CACAGCCTTGGTGGACCCCAAAAGAAACGGGCTACAGAGTTAGTAAGTCTTTAGAAA 434

Qy 1786 GAACTGGAAGTTAATCCGAGATGAAGCCTTGGAGTATGATGATGAATGAAGGCTCTCT 1845

Db 435 GAACTGGAAGTTAATCCGAGATGAAGCCTTGGAGTATGATGATGAATGAAGGCTCTCT 494

Qy 1846 TCCTGCTGAGATGAAAAACCTTGGAGGAAAAAGGGGACTTGGAGCCAGTTCAGCTGTGGC 1905

Db 495 TCCTGCTGAGATGAAAAACCTTGGAGGAAAAAGGGGACTTGGAGCCAGTTCAGCTGTGGC 554

Qy 1906 AGCAAGGAAGAAGA 1919

Db 555 AGCAAGGAAGAAGA 568

RESULT 27

BQ889709 857 bp mRNA linear EST 16-AUG-2002

LOCUS BQ889709

DEFINITION AGENCOURT 8118183 Lupski dorsal root ganglion Homo sapiens cDNA clone IMAGE:6181443 5', mRNA sequence.

ACCESSION BQ889709.1 GI:22281723

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 857)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgaabs-x@mail.nih.gov Tissue Procurement: Dr. James R. Lupski CDNA Library Preparation: Life Technologies, Inc. DNA sequencing by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13565 row: n column: 04 High quality sequence stop: 590.

FEATURES

Location/Qualifiers

1..857

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6181443"

/sex="male"

/tissue\_type="dorsal root ganglia"

/dev\_stage="adult, 36 yr"

/lab\_host="DH10B"

/clone\_lib="Lupski dorsal root ganglion"

/note=vector: pCMV-SPORT6 (Life Technologies); Site\_1: NotI; Site\_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACCGCTCCG-3' and 5'-GACTAGTCTTAGTCGGAGCGGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 19.4%; Score 450; DB 5; Length 857;

Best Local Similarity 99.8%; Pred. No. 9.8e-230; Indels 0; Gaps 0;

Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCATATGGCCAGCGTGAAGATGCCAAGAGCAGCGCAACAGCAGCAGCG 60





cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2139 row: i column: 04  
High quality sequence stop: 595.

Location/Qualifiers  
1. .1014  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5952075"  
/tissue\_type="ductal carcinoma, cell line"  
/lab\_host="NIH\_MGC\_110"  
/clone\_lib="NIH\_MGC\_110"  
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 19.4%; Score 450; DB 5; Length 1014;  
Best Local Similarity 99.8%; Pred. No. 1e-229;  
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACCGTGAATGCGCCAGCGTGAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 103 CGGACCGTGAATGCGCCAGCGTGAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 162  
QY 61 GCTCCGCGCAGCGGTAGCAGAGTGGCGCAGCAGCAGCGCCGCGGCGCCGAGAGAGACAA 120  
DB 163 GCTCCGCGCAGCGGTAGCAGAGTGGCGCAGCAGCAGCGCCGCGGCGCCGAGAGAGACAA 222  
QY 121 AGCATGGAGGACACAGAGATGGGAGAAAGCGGACTCTCGGGAACCTTCACTTTCACGT 180  
DB 223 AGCATGGAGGACACAGAGATGGGAGAAAGCGGACTCTCGGGAACCTTCACTTTCACGT 282  
QY 181 GGTATTATGTTGATTCGATTCGTCGGCGTCTGACATCTGTAGCTGTGCTTTGGTTTGCATC 240  
DB 283 GGTATTATGTTGATTCGATTCGTCGGCGTCTGACATCTGTAGCTGTGCTTTGGTTTGCATC 342  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAACTATGATGCTGATGTTGATGGAG 300  
DB 343 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAACTATGATGCTGATGTTGATGGAG 402  
QY 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
DB 403 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 462  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGAGAGGTTCTCTGTGG 420  
DB 463 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGAGAGGTTCTCTGTGG 522  
QY 421 AGGCAGAACCCCGAGATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 480  
DB 523 AGGCAGAACCCCGAGATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 582  
QY 481 AAATGGTACACGCAACATG 501  
DB 583 AAATGGTACACGCAACATG 603

RESULT 30  
AUI32666  
LOCUS  
DEFINITION NT2RP4 Homo sapiens cDNA clone NT2RP4000310 5', mRNA  
sequence.  
AUI32666  
ACCESSION

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AUI32666.1 GI:10993205  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS

1 (bases 1 to 767)  
Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y., Saito, K., Yamamoto, J.,  
Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuho, Y. and  
Isogai, T.  
HRI human cDNA project (Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y.,  
Saito, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T.,  
Sugano, S., Masuho, Y., Isogai, T.)  
Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5' & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

TITLE  
JOURNAL  
COMMENT

Location/Qualifiers  
1. .767  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP4000310"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/clone\_lib="NT2RP4"  
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"

FEATURES  
source

Location/Qualifiers  
1. .767  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP4000310"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/clone\_lib="NT2RP4"  
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Query Match 19.3%; Score 448; DB 1; Length 767;  
Best Local Similarity 99.6%; Pred. No. 1.e-228;  
Matches 548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGACCGTGAATGCGCCAGCGTGAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 181 CGGACCGTGAATGCGCCAGCGTGAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 240  
QY 61 GCTCCGCGCAGCGGTAGCAGAGTGGCGCAGCAGCAGCGCCGCGGCGCCGAGAGAGACAA 120  
DB 241 GCTCCGCGCAGCGGTAGCAGAGTGGCGCAGCAGCAGCGCCGCGGCGCCGAGAGAGACAA 300  
QY 121 AGCATGGAGGACACAGAGATGGGAGAAAGCGGACTCTCGGGAACCTTCACTTTCACGT 180  
DB 301 AGCATGGAGGACACAGAGATGGGAGAAAGCGGACTCTCGGGAACCTTCACTTTCACGT 360  
QY 181 GGTATTATGTTGATTCGATTCGTCGGCGTCTGGAACATCTGTAGCTGTGCTTTGGTTTGCATC 240  
DB 361 GGTATTATGTTGATTCGATTCGTCGGCGTCTGGAACATCTGTAGCTGTGCTTTGGTTTGCATC 420  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAACTATGATGCTGATGTTGATGGAG 300  
DB 421 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAACTATGATGCTGATGTTGATGGAG 480  
QY 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
DB 481 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGAGAGGTTCTCTGTGG 420  
DB 541 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGAGAGGTTCTCTGTGG 600  
QY 421 AGGCAGAACCCCGAGATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 480  
DB 601 AGGCAGAACCCCGAGATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 660





**TITLE** Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation

**JOURNAL** Nat. Biotechnol. 22 (6), 707-716 (2004)

**COMMENT** Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 695 Std Error: 0.00.  
Location/Qualifiers  
1..695  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/note="Oligo dT primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

# FEATURES

source

## ORIGIN

Query Match 18.5%; Score 430; DB 7; Length 695;  
Best Local Similarity 99.8%; Pred. No. 5.4e-219;  
Matches 480; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATCGCCAGCGTAAGAATGCCAAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 215 CGGACCGTGCATCGCCAGCGTAAGAATGCCAAGCAGCGGCAACAGCAGCAGCAGCG 274  
QY 61 GCTCCGGCAGCGGTAGCAGAGTCGGCGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
DB 275 GCTCCGGCAGCGGTAGCAGAGTCGGCGCAGCAGCGCCCGGGCCCGGAGAGACAA 334  
QY 121 AGCATGGAGGACACAGAATGGAGAAAGCGCGACTCTCGGGAACCTTCATCTTCACGT 180  
DB 335 AGCATGGAGGACACAGAATGGAGAAAGCGCGACTCTCGGGAACCTTCATCTTCACGT 394  
QY 181 GGTATTATGGTGAATTCCTGGCGCTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC 240  
DB 395 GGTATTATGGTGAATTCCTGGCGCTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC 454  
QY 241 TTGTTGACTATGAGGAATTCCTAGGAAACTAGGAATCTATGATGCTGTGATGGAG 300  
DB 455 TTGTTGACTATGAGGAATTCCTAGGAAACTAGGAATCTATGATGCTGTGATGGAG 514  
QY 301 ATTTTGTATGGATGATGCCAAAGTTTATTAGGACTTAAGAGAGATCTACTTCAGAGC 360  
DB 515 ATTTTGTATGGATGATGCCAAAGTTTATTAGGACTTAAGAGAGATCTACTTCAGAGC 574  
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGGTTCTCTGTGG 420  
DB 575 CAGCAGTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGGTTCTCTGTGG 634  
QY 421 AGGCAGAACCCAGAAATATCGAAGATGAAGCAAGAAGCAAAATTCAGTCCCTTCTCCATG 480  
DB 635 AGGCAGAACCCAGAAATATCGAAGATGAAGCAAGAAGCAAAATTCAGTCCCTTCTCCATG 694  
QY 481 A 481  
DB 695 A 695

RESULT 35  
AL700393  
LOCUS DKEZp686K01118 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
DEFINITION DKEZp686K01118 5', mRNA sequence.  
ACCESSION AL700393  
VERSION AL700393.1 GI:19620926

# KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann,S.).  
EST (Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and  
Wiemann,S.)  
Unpublished (1999)  
Contact: MIPS  
MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No s1 sequence available.  
This clone (DKFZp686K01118) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

source  
1..554  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686K01118"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hlcc3)"  
/notes="Vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"

## ORIGIN

Query Match 18.3%; Score 425; DB 1; Length 554;  
Best Local Similarity 99.8%; Pred. No. 2.6e-216;  
Matches 475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1439 TTATGAAGAGGTCTCAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTT 1498  
DB 1 TTATGAAGAGGTCTCAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTT 60  
QY 1499 CATCTGACGACACAGAAATTCCTGAGAGCATCCCATATTTAAAGAGGAATAGA 1558  
DB 61 CATCTGACGACACAGAAATTCCTGAGAGCATCCCATATTTAAAGAGGAATAGA 120  
QY 1559 ATCCGAGATCTTGGCACTGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCA 1618  
DB 121 ATCCGAGATCTTGGCACTGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCA 180  
QY 1619 GAGGTTTGGGAAACAAAGAGGCATATAAGTGTGATGAGCTTTGGGCACACAGAGAGCACATT 1678  
DB 181 GAGGTTTGGGAAACAAAGAGGCATATAAGTGTGATGAGCTTTGGGCACACAGAGAGCACATT 240  
QY 1679 TGCACTGTCTGGCAACGCTCACTCTCAATGTAATGAGTGAAGACACAGCTTGGTG 1738  
DB 241 TGCACTGTCTGGCAACGCTCACTCTCAATGTAATGAGTGAAGACACAGCTTGGTG 300  
QY 1739 GACCCCAAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACCTGAAGTT 1798  
DB 301 GACCCCAAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACCTGAAGTT 360  
QY 1799 AATCCGAGATGAAGGCTTTGCACTGATGATAAGCCAAAGGCTCTCTTCTCCCTGAGGA 1858  
DB 361 AATCCGAGATGAAGGCTTTGCACTGATGATAAGCCAAAGGCTCTCTTCTCCCTGAGGA 420  
QY 1859 TGAACACCTGAGGAAAAAGGGGACTTGAGCAGTTCAGCTGTGTGACAGCAAGAA 1914  
DB 421 TGAACACCTGAGGAAAAAGGGGACTTGAGCAGTTCAGCTGTGTGACAGCAAGAA 476

```
RESULT 36
BU502544      959 bp      mRNA      linear      EST 12-SEP-2002
LOCUS      AGENCOURT_10023079 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6480842
DEFINITION      5', mRNA sequence.
ACCESSION      BU502544
VERSION      BU502544.1 GI:22807433
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2661 row: e column: 03
High quality sequence stop: 591.
FEATURES      Location/Qualifiers
source      1..959
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6480842"
            /tissue_type="carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_40"
            /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
            Site 2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."
ORIGIN
Query Match      18.3%; Score 425; DB 5; Length 959;
Best Local Similarity 99.4%; Pred. No. 2.8e-216; Mismatches 4; Indels 0; Gaps 0;
Matches 625; Conservative 0;

QY      1      CGGACCGTGCATGGCCCGGCGTAAAGATGCCAAGAGCAGCGGCGCAACAGCAGCAGCG 60
DB      40     CGGACCGTGCATGGCCCGGCGTAAAGATGCCAAGAGCAGCGGCGCAACAGCAGCAGCG 99

QY      61     GTCGCGGAGCAGCGTACGAGTGGCGGCGAGCAGAGCCCGGCGCGGAGAGACAA 120
DB      100    GCTCCGGCAGCCGTAGCAGTGGCGGCGAGCAGAGCCCGGCGCGGAGAGACAA 159

QY      121    AGCATGGAGGACACAGAGATGGAGAAAGCGGACTCTCGGGAACCTTCATCTTCACGT 180
DB      160    AGCATGGAGGACACAGAGATGGAGAAAGCGGACTCTCAGAACTTCATCTTCACGT 219

QY      181    GGTATTATGTTGATTCGATTCGTCGGCGCTCGACATCTGTAGCTGCGTTGTTGATC 240
DB      220    GGTATTATGTTGATTCGATTCGTCGGCGCTCTGAACATCTGTAGCTGCGTTGTTGATC 279

QY      241    TTGTTGACTATAGGGAAGTTCTAGGAAACTAGGATCTATCATCTGCTGATGCTGATGGAG 300
DB      280    TTGTTGACTATAGGGAAGTTCTAGGAAACTAGGATCTATCATCTGCTGATGCTGATGGAG 339

QY      301    ATTTTTCATGTTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360
DB      340    ATTTTTCATGTTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC 399
```

```
QY      361    CAGCAGTCCCGCAGAGAGAGGCTGAGCCACACACTGAGCCCGCAGGAGCAGGTTCTCTGG 420
DB      400    CAGCAGTCCCGCAGAGAGAGGCTGAGCCACACACTGAGCCCGCAGGAGCAGGTTCTCTGG 459

QY      421    AGGCAGACCCCGAGATATCGAAGATGAGCAAAAGAACAAATTTCAGTCCCTTCTCCATG 480
DB      460    AGGCAGACCCCGAGATATCGAAGATGAGCAAAAGAACAAATTTCAGTCCCTTCTCCATG 519

QY      481    AAATGTGACACGAGAACATGTTGAGGGGAGAGACTTGCACAAAGAGATGACCCACAG 540
DB      520    AAATGTGACACGAGAACATGTTGAGGGGAGAGACTTGCACAAAGAGATGACCCACAG 579

QY      541    GAGAACCCACAAAGAGAGATGATGATTCTTATGGCGACTGATGTAGATGATGATTTG 600
DB      580    GAGAACCCACAAAGAGAGATGATGATTCTTATGGCGACTGATGTAGATGATGATTTG 639

QY      601    AGACCCCTGGAACTGAAATGATCTCATGAA 629
DB      640    AGACCCCTGGAACTGAAATGATCTCATGAA 668

RESULT 37
AL120973      424 bp      mRNA      linear      EST 04-SEP-2003
LOCUS      DKFZP762H093_r1 762 (synonym: hm12) Homo sapiens cDNA clone
DEFINITION      DKFZP762H093 5', mRNA sequence.
ACCESSION      AL120973
VERSION      AL120973.1 GI:5926974
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 424)
AUTHORS      Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE      EST (Koehler, et al.)
JOURNAL      Unpublished (1999)
COMMENT      Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No s1 sequence available.
            This clone (DKFZP762H093) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES      Location/Qualifiers
source      1..424
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZP762H093"
            /tissue_type="melanoma (MeWo cell line)"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="762 (synonym: hm12)"
            /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN
Query Match      18.2%; Score 424; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 8.4e-216; Mismatches 0; Indels 0; Gaps 0;
Matches 424; Conservative 0;

QY      264    GGAAACTAGGAATCTATGATCGTGTGATCGATCGAGATTTTGATGTGGATGATGCCAAA 323
DB      1      GGAAACTAGGAATCTATGATCGTGTGATCGATCGAGATTTTGATGTGGATGATGCCAAA 60

QY      324    GTTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGCCAGCAGTCCCGCCAGAGAGGCT 383
```

Db 61 GTTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCGCCAGAGAGGCT 120  
 Qy 384 GAGCCACACACTGAGCCGAGGAGCAGTCTCTGTGGAGGAGAGACCCAGCAATATCGAA 443  
 Db 121 GAGCCACACACTGAGCCGAGGAGCAGTCTCTGTGGAGGAGAGACCCAGCAATATCGAA 180  
 Qy 444 GATGAGCAAAAGAAACAAATTCAGTCCCTCTCCATGAATGTTACAGCAGACATGTT 503  
 Db 181 GATGAGCAAAAGAAACAAATTCAGTCCCTCTCCATGAATGTTACAGCAGACATGTT 240  
 Qy 504 GAGGAGAGAGCTTGCACAGAGAGATGGACCCACAGAGAGAACCAACAGAGAGATGAT 563  
 Db 241 GAGGAGAGAGCTTGCACAGAGAGATGGACCCACAGAGAGAACCAACAGAGAGATGAT 300  
 Qy 564 GAGTTTCTTATGGGAGCTGATGTAGATGATAGATTGAGACCTGGAACTGAGTATCT 623  
 Db 301 GAGTTTCTTATGGGAGCTGATGTAGATGATAGATTGAGACCTGGAACTGAGTATCT 360  
 Qy 624 CATGAAGAAACCGAGCATAGTTACCGTGGAGAGACAGATTTCACAGAGCTGTAATCAG 683  
 Db 361 CATGAAGAAACCGAGCATAGTTACCGTGGAGAGACAGATTTCACAGAGCTGTAATCAG 420  
 Qy 684 GATA 687  
 Db 421 GATA 424

## RESULT 38

LOCUS CN481470  
 DEFINITION hw08b05.y1 Human primary human ocular pericytes. Unamplified (hw)  
 ACCESSION Homo sapiens cDNA clone hw08b05 5', mRNA sequence.  
 VERSION CN481470  
 KEYWORDS EST.  
 SOURCE CN481470.1 GI:46562974

## ORGANISM

Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Tsai, J.Y. and Wistow, G.

Expressed sequence tag analysis of cultured primary human ocular pericytes

## JOURNAL

Unpublished (2004)  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 08 row: b column: 05

Seq primer: M13RP1 reverse primer (ABI).

Location/Qualifiers

## FEATURES

source

1..606  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="hw08b05"  
 /cell\_type="pericytes"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human primary human ocular pericytes."  
 /unamplified (hw)

/note="Organ: Eye; Vector: pSport1; RNA was extracted from primary human pericytes in culture. A directionally cloned cDNA library in the pSPOR1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was

carried out using a Not I primer-adaptor  
 [5'-pGATAGTCTAGATCGGAGCGCCGCTT15-3']. cDNA was  
 cloned in Not I/Sal I sites. EST analysis was performed at  
 the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Query Match 18.2%; Score 424; DB 7; Length 606;  
 Best Local Similarity 99.8%; Pred. No. 9e-216;  
 Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 CGGACCGTGCATTAATGGCCAGCGTAAGATGCGGAGCAGCGGCAACAGCAGCAGCGG 60  
 Db 132 CGGACCGTGCATTAATGGCCAGCGTAAGATGCGGAGCAGCGGCAACAGCAGCAGCGG 191  
 Qy 61 GCTCCGGCAGCGGTAGCAGCAGTGGCGGAGCAGCGGCGGCGGCGGCGGCGGAGAGACAA 120  
 Db 192 GCTCCGGCAGCGGTAGCAGCAGTGGCGGAGCAGCGGCGGCGGCGGCGGCGGAGAGACAA 251  
 Qy 121 AGCATGGAGGACACAGATGGAGAGGAGGAGGAGCTCTCGGGAATCTTCACTTTCACGT 180  
 Db 252 AGCATGGAGGACACAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 311  
 Qy 181 GGTATTGGTGATGTCATTGCTGGCGCTCTGAGACATCTGTAGCTGTGTTGTTGATC 240  
 Db 312 GGTATTGGTGATGTCATTGCTGGCGCTCTGAGACATCTGTAGCTGTGTTGTTGATC 371  
 Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGAG 300  
 Db 372 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGAG 431  
 Qy 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
 Db 432 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 491  
 Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGAGGAGGAGG 420  
 Db 492 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGAGGAGGAGG 551  
 Qy 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCT 475  
 Db 552 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCT 606

## RESULT 39

LOCUS BF943222  
 DEFINITION QV2-NN0045-041000-403-b08 NN0045 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF943222  
 VERSION BF943222.1 GI:12360497  
 KEYWORDS EST.

## SOURCE

Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 504)

## REFERENCE

AUTHORS  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-NN0045-041000-403-b08&t3=2000-10-04&t4=1)  
041000-403-b08&t3=2000-10-04&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 504.

FEATURES  
Location/Qualifiers  
1..504  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="NN0045"  
/note="Organ: nervous normal; Vector: puc18; Site: 1: Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196, 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN  
Query Match 18.2%; Score 423; DB 4; Length 504;  
Best Local Similarity 99.8%; Pred. No. 3e-215;  
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 332 AGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 391  
DB |||||  
31 AGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 90  
QY 392 CACTGAGCCGAGGAGCAGTCTCTGTGGAGGCGAGACCCAGAAATATCGAAGATGAAGC 451  
DB |||||  
91 CACTGAGCCGAGGAGCAGTCTCTGTGGAGGCGAGACCCAGAAATATCGAAGATGAAGC 150  
QY 452 AAACGAAACAATTCAGTCCCTTCTCCATGAATGTGTACACGACAGAAATGTGTGAGGAGA 511  
DB |||||  
151 AAAGAGACAAATTCAGTCCCTTCTCCATGAATGTGTACACGACAGAAATGTGTGAGGAGA 210  
QY 512 AGACTTGCNACAAAGAGATGAGCCACAGGAGAACCAACAGAGGATGATGATTTCT 571  
DB |||||  
211 AGACTTGCNACAAAGAGATGAGCCACAGGAGAACCAACAGAGGATGATGATTTCT 270  
QY 572 TATGGGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 631  
DB |||||  
271 TATGGGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330  
QY 632 AACCCGACATAGTTACCGCTGGAAGAGACAGTTTCAAGACTGTATCAGGATATGGA 691  
DB |||||  
331 AACCCGACATAGTTACCGCTGGAAGAGACAGTTTCAAGACTGTATCAGGATATGGA 390  
QY 692 AGAGATGATGCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAAGATGAAG 751  
DB |||||  
391 AGAGATGATGCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAAGATGAAG 450  
QY 752 ATTGCACCATGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 805  
DB |||||  
451 ATTGCACCATGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 504

RESULT 40  
CN481475  
LOCUS  
DEFINITION  
hw08b10.y1 Human primary human ocular pericytes. Unamplified (hw)  
Homo sapiens cDNA clone hw08b10 5', mRNA sequence.  
CN481475  
ACCESSION  
VERSION  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 610)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Tsai, J.Y. and Wistow, G.  
Expressed sequence tag analysis of cultured primary human ocular pericytes  
Unpublished (2004)  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 08 row: b column: 10  
Seq primer: M13RP1 reverse primer (ABI).  
Location/Qualifiers  
1..610  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="hw08b10"  
/cell\_type="pericytes"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human primary human ocular pericytes."  
Unamplified (hw)  
/note="Organ: Eye; Vector: pSport1; RNA was extracted from primary human pericytes in culture. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTTCTAGATCGGCGGCCCT(T)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 18.2%; Score 423; DB 7; Length 610;  
Best Local Similarity 99.8%; Pred. No. 3.1e-215;  
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACCGTGAATGGCCGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB |||||  
137 CGGACCGTGAATGGCCGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 196  
QY 61 GCTCCGCGCAGCGGTAGCAGAGTCCGGGCGAGCAGCGCCCGGGCGCCGAGAGAGACAA 120  
DB |||||  
197 GCTCCGCGCAGCGGTAGCAGAGTCCGGGCGAGCAGCGCCCGGGCGCCGAGAGAGACAA 256  
QY 121 AGCATGAGGACACAGAAATGGGAGGAAAGCGGACTCTCGGAACTTCACTTCTCAGCT 180  
DB |||||  
257 AGCATGAGGACACAGAAATGGGAGGAAAGCGGACTCTCAGGAACTTCACTTCTCAGCT 316  
QY 181 GGTATTATGTTGATGTCATTGCTGGCGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 240  
DB |||||  
317 GGTATTATGTTGATGTCATTGCTGGCGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 376  
QY 241 TTGTTGACTATGAGGAAAGTTCTAGGAAATCTAGGAATCTATGATCTGATGTTGATGAG 300  
DB |||||  
377 TTGTTGACTATGAGGAAAGTTCTAGGAAATCTAGGAATCTATGATCTGATGTTGATGAG 436  
QY 301 ATTTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 360  
DB |||||  
437 ATTTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 496  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGTTCTCTGTGG 420  
DB |||||  
497 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGTTCTCTGTGG 556  
QY 421 AGGAGAAACCCCGAGAAATTCGAAGATGAGCAAGAAAGACAAATTCAGTCCCTTC 474  
DB |||||  
557 AGGAGAAACCCCGAGAAATTCGAAGATGAGCAAGAAAGACAAATTCAGTCCCTTC 610

```

RESULT 41
AUI38660
LOCUS
DEFINITION
AUI38660 PLACE1 Homo sapiens cDNA clone PLACE1009036 5', mRNA
sequence.
ACCESSION
AUI38660
VERSION
AUI38660.1 GI:11000181
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 790)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Ishigai,T.
HRI human cDNA project
Unpublished (2000)
JOURNAL
Genomics Laboratory
COMMENT
Contact: Takao Isogai
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..790
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1009036"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="Vector: pME18SFL3"
ORIGIN
Query Match 18.1%; Score 420; DB 1; Length 790;
Best Local Similarity 100.0%; Pred. No. 1.3e-213;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 749 AAGATTGCACCATGATACAGATGATGTAACATACCAAGTCTATGAGGAACAAGCAGTATA 808
DB 1 AAGATTGCACCATGATACAGATGATGTAACATACCAAGTCTATGAGGAACAAGCAGTATA 60
QY 809 TGAACCTCTAGAAAATGAAGGGATAGAATACAGAGTAAGTCTCCCTCGAGGATAA 868
DB 61 TGAACCTCTAGAAAATGAAGGGATAGAATACAGAGTAAGTCTCCCTCGAGGATAA 120
QY 869 TCCTGTAGAGATTCACAGTAATGTTAGAGAGTAAGATTTTCCTGTGGAGAGACA 928
DB 121 TCCTGTAGAGATTCACAGTAATGTTAGAGAGTAAGATTTTCCTGTGGAGAGACA 180
QY 929 GCAGAGATTCACAGAGAAACAAATAGAAAAACAGATGATCCAGAAACAAAAAGT 988
DB 181 GCAGAGATTCACAGAGAAACAAATAGAAAAACAGATGATCCAGAAACAAAAAGT 240
QY 989 TAAGAAAAAGAGCCCTAAACTTTTAAATAAATTTGATAAGACTATTAAAGCTGAACCTGA 1048
DB 241 TAAGAAAAAGAGCCCTAAACTTTTAAATAAATTTGATAAGACTATTAAAGCTGAACCTGA 300
QY 1049 TGCTGCAGAAAAACTCCGTAAAGGGGAAAAATTTGAGGAGAGCTGAATGCAATTTAAAGA 1108
DB 301 TGCTGCAGAAAAACTCCGTAAAGGGGAAAAATTTGAGGAGAGCTGAATGCAATTTAAAGA 360
QY 1109 ACTAGTACGGAATACCTCTAGAGTCCACGAGCAAGATATGGGAAGGCGCAGTGTGAGGA 1168
DB 361 ACTAGTACGGAATACCTCTAGAGTCCACGAGCAAGATATGGGAAGGCGCAGTGTGAGGA 420

```

---

```

RESULT 42
BQ318605
LOCUS
DEFINITION
CM0-CT0549-210900-561-a04 CT0549 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BQ318605
VERSION
BQ318605.1 GI:20924374
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 433)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL
MEDLINE
20202663
PUBMED
10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM0&t2=CM0-CT0549-
210900-561-a04&t3=2000-09-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 433.
Location/Qualifiers
1..433
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0549"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
ORIGIN
Query Match 18.0%; Score 418; DB 5; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.4e-212;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 479 TGAATGGTTACACGACGACACATCTTGAGGGAGAGACTTGCACACAGAGATGGACCCAC 538
DB 16 TGAATGGTTACACGACGACACATCTTGAGGGAGAGACTTGCACACAGAGATGGACCCAC 75
QY 539 AGAGAACACCAACCAAGAGGATGATGTTCTTATGGCGACTGATGTAGATGATGATT 598
DB 76 AGAGAACACCAACCAAGAGGATGATGTTCTTATGGCGACTGATGTAGATGATT 135
QY 599 TGAGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTACCACGCTGGAGA 658
DB 136 TGAGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTACCACGCTGGAGA 195
QY 659 GACAGTTTCCAAAGACTCTAATCAGGATATGGAAGAGATGATCTGAGCAGGAAATCC 718

```



```
Db 196 GACAGTTTCACAGACTGTAAATCAGGATATGGAAGAGATGATCTCTGAGCAGCAAAAATCC 255
Qy 719 AGATTCCAGTGAACCACTAGTAGAAGATGAAGATTGCACCATGATACAGATGATTAAC 778
Db 256 AGATTCCAGTGAACCACTAGTAGAAGATGAAGATTGCACCATGATACAGATGATTAAC 315
Qy 779 ATACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAAT 838
Db 316 ATACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAAT 375
Qy 839 CACAGAAGTAACTCTCTCCCTCAGGATTAATCTCTGTAGAAGATTCACAGTAAATTGTA 896
Db 376 CACAGAAGTAACTCTCTCCCTCAGGATTAATCTCTGTAGAAGATTCACAGTAAATTGTA 433

RESULT 43
LOCUS BQ278505 1100 bp mRNA linear EST 07-MAY-2002
DEFINITION AGENCOURT_6773215 NIH_MGC_128 Homo sapiens cDNA clone IMAGE:5811567
5', mRNA sequence.
ACCESSION BQ278505
VERSION BQ278505.1 GI:20488713
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2060 row: j column: 16
High quality sequence stop: 394.
Location/Qualifiers
1..1100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5811567"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_128"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggccgcttcggcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
salivary gland - 1.3%, and skin - 2.3%). 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCTGTATACAGCAGATGGCCATTCAGCGGG-3' and
5'-ATTCTAGAGCCGCGGCGGCGCATG-AT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the >2 kb
size fraction (other fractions present in NIH_MGC_126 and
NIH_MGC_127). Library created in the laboratory of T.
Usdin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH_MGC
Library."

FEATURES
source
17.9%; Score 417; DB 5; Length 1100;
Query Match 100.0%; Pred. No. 5.9e-212; Indels 0; Gaps 0;
Best Local Similarity 0; Mismatches 0;
Matches 417; Conservative 0;

196 GACAGTTTCACAGACTGTAAATCAGGATATGGAAGAGATGATCTCTGAGCAGCAAAAATCC 335
115 ATCTATGATGCTGATGGTGGAGATTTCATGATGGATGATGCCAAAAGTTTATTAGGA 174
336 CTTAAAGAGAGATCTACTCTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACT 395
175 CTTAAAGAGAGATCTACTCTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACT 234
396 GAGCCCGAGGAGCAGGTTCTCTGTGAGGCAGAAACCCAGAAATATCGAAGATGAACAAAA 455
235 GAGCCCGAGGAGCAGGTTCTCTGTGAGGCAGAAACCCAGAAATATCGAAGATGAACAAAA 294
456 GAACAAATTCAGTCCCTCTCCATGAAATGGTACGCGAGAACATGTTGAGGGAAGAGAC 354
516 TTGCAACAAGAGATGGAGCCACAGAGAACACACAAAGAGAGATGATGAGTTCTTATG 575
355 TTGCAACAAGAGATGGAGCCACAGAGAACACACAAAGAGAGATGATGAGTTCTTATG 414
576 GCGACTGATGTAGATGATAGATTTCAGACCCCTGGAACTGAAAGTATCTCATGAAGAAC 635
415 GCGACTGATGTAGATGATAGATTTCAGACCCCTGGAACTGAAAGTATCTCATGAAGAAC 474
636 GAGCATAGTTACACGTGGAAGAGACAGTTTCACAAGACTGTAATCAGCATATGGAA 692
475 GAGCATAGTTACACGTGGAAGAGACAGTTTCACAAGACTGTAATCAGCATATGGAA 531

RESULT 44
LOCUS BF962946 444 bp mRNA linear EST 22-JAN-2001
DEFINITION QV2-NN0045-181200-563-all NN0045 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF962946
VERSION BF962946.1 GI:12380221
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 444)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-NN0045-
181200-563-all&t3=2000-12-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 444.
Location/Qualifiers
1..444
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
```

/clone\_lib="NN0045"  
 /note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent Application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Query Match 17.9%; Score 415; DB 4; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-211;  
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 391  
 DB 30 AGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 89  
 QY 392 CACTGAGCCCGAGGAGCAGGTTCCTGTGGAGGAGAGCCAGAGATATCGAAGATGAAGC 451  
 DB 90 CACTGAGCCCGAGGAGCAGGTTCCTGTGGAGGAGAGCCAGAGATATCGAAGATGAAGC 149  
 QY 452 AAAAGAGACAAATTCAGTCCCTCTCCATGAAATGGTACACGAGAACATGTTGAGGAGA 511  
 DB 150 AAAGAGACAAATTCAGTCCCTCTCCATGAAATGGTACACGAGAACATGTTGAGGAGA 209  
 QY 512 AGACTTGCACAGAGATGAGCCAGAGAGAACCAACAGAGGATGATGATTTCT 571  
 DB 210 AGACTTGCACAGAGATGAGCCAGAGAGAACCAACAGAGGATGATGATTTCT 269  
 QY 572 TATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 631  
 DB 270 TATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 329  
 QY 632 AACCCAGATATTTACCACTGGAAGAGACAGTTTCAAGACTGTAATCAGATATGGA 691  
 DB 330 AACCCAGATATTTACCACTGGAAGAGACAGTTTCAAGACTGTAATCAGATATGGA 389  
 QY 692 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 746  
 DB 390 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 444

## RESULT 45

BQ639558  
 LOCUS  
 DEFINITION  
 hel7b04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he  
 Homo sapiens cDNA clone hel7b04 5', mRNA sequence.  
 BQ639558  
 VERSION  
 BQ639558.1 GI:21764017  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 570)  
 Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,  
 Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.  
 Expressed sequence tag analysis of human retina for the NEIBank  
 Project: Retbindin, an abundant, novel retinal cDNA and alternative  
 splicing of other retina-preferred gene transcripts  
 Mol. Vis. 8 (4), 196-204 (2002)  
 22103461  
 12107411  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 17 row: b column: 04  
 Seq primer: M13RP1 reverse primer (ABI).

## JOURNAL

MEDLINE  
 PUBMED  
 COMMENT

## FEATURES

## source

Location/Qualifiers  
 1..570  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="hel7b04"  
 /tissue\_type="Retina"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Retina cDNA (Un-normalized,  
 unamplified): hd/he"  
 /note="Organ: Eye; Vector: pSPOR1; Neural retina tissue  
 was dissected from two 80 year old donors with no observed  
 eye disease. 100ug of total RNA was used for library  
 construction. A directionally cloned cDNA library in the  
 pSPOR1 vector (Life Technologies) was constructed at  
 Bioberve Biotechnology (Laurel MD) essentially following  
 the protocols of the SuperScript plasmid System full  
 details of which are contained in the manufacturer's  
 instruction manual (http://www.lifetech.com/). First  
 strand synthesis was carried out using a Not I  
 primer-adaptor  
 [5'-pGACTAGTTCTAGATCGCGGCGCCGCT(15-3')]. EST analysis  
 was performed on the unamplified library at the NIH  
 Intramural Sequencing Center (NISC)."

## ORIGIN

Query Match 17.6%; Score 409; DB 5; Length 570;  
 Best Local Similarity 100.0%; Pred. No. 1e-207;  
 Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1628 GAACAAGAGGCGATATAAGTGTATGAGTTGGGCAAGAGAGGACACTTTGCATCTGT 1687  
 DB 162 GAACAAGAGGCGATATAAGTGTATGAGTTGGGCAAGAGAGGACACTTTGCATCTGT 221  
 QY 1698 CTGGCAACGCTCACTCTACAAATGTGAATGGACTGAAAGACAGAGCTTTGGTGGACCCCAA 1747  
 DB 222 CTGGCAACGCTCACTCTACAAATGTGAATGGACTGAAAGACAGAGCTTTGGTGGACCCCAA 281  
 QY 1748 AGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAGTTAATCCGAGA 1807  
 DB 282 AGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAGTTAATCCGAGA 341  
 QY 1808 TGAAGGCTTGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1867  
 DB 342 TGAAGGCTTGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 401  
 QY 1868 GAGGAAAAAGGGGACTGGAGCCAGTTTCAAGTGTGGCAGCAAGAGGAAGAAATGAAAA 1927  
 DB 402 GAGGAAAAAGGGGACTGGAGCCAGTTTCAAGTGTGGCAGCAAGAGGAAGAAATGAAAA 461  
 QY 1928 TGCCTGCAAGAGGAGTCTCTTAAACCTGTACCTTACTAGAAAGTTCCTCCGAGACACAGG 1987  
 DB 462 TGCCTGCAAGAGGAGTCTCTTAAACCTGTACCTTACTAGAAAGTTCCTCCGAGACACAGG 521  
 QY 1988 ATGCAGAGAGGAGACAGATCAATATTCATGACGACCCCGGAGCTCAC 2036  
 DB 522 ATGCAGAGAGGAGACAGATCAATATTCATGACGACCCCGGAGCTCAC 570

## RESULT 46

BQ639558  
 LOCUS  
 DEFINITION  
 1700600180518 GRN\_PRENU Homo sapiens cDNA 5', mRNA sequence.  
 BQ639558  
 ACCESSION  
 BQ639558.1 GI:47368755  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 740)  
 Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,

Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,  
 Lebkowski,J and Stanton,L.W.  
 Transcriptional characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 CONTACT: Brandenberger R  
 Regenerative Medicine  
 Genon Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@genon.com  
 Insert Length: 740 Std Error: 0.00.

FEATURES  
 source  
 1. 740  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cell, retinoic acid and  
 mitogen-treated HES cell line H7"  
 /clone\_lib="GRN\_PREENU"  
 /note="Oligo dT primed, full-length enriched cDNA library  
 from HES cell line H7 (p29) maintained in feeder-free  
 conditions. Embryoid bodies were generated in the presence  
 of all-trans retinoic acid and mitogens."

ORIGIN  
 Query Match 17.3%; Score 402; DB 7; Length 740;  
 Best Local Similarity 99.8%; Pred. NO. 6.4e-204;  
 Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 49 GCAGCAGACGGCTCCGGCAGCGGTACGACGAGTGGCGCAGCAGCAGCGCCCGGGGCC 108  
 Db 1 GCAGCAGACGGCTCCGGCAGCGGTACGACGAGTGGCGCAGCAGCAGCGCCCGGGGCC 60  
 QY 109 GGAGAGACAAAGCATGGAGACACAAAGATGGAGGAGAAAGCGGACTCTCGGGAATT 168  
 Db 61 GGAGAGACAAAGCATGGAGACACAAAGATGGAGGAGAAAGCGGACTCTCGGGAATT 120  
 QY 169 CATCTTCACGTGGTTATGTTGATGATGCTGCTGGCGTCTGGACATCTCTAGCTGTCG 228  
 Db 121 CATCTTCACGTGGTTATGTTGATGATGCTGCTGGCGTCTGGACATCTCTAGCTGTCG 180  
 QY 229 TTTGGTTTCATCTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTG 288  
 Db 181 TTTGGTTTCATCTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTG 240  
 QY 289 ATGGTGATGGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 348  
 Db 241 ATGGTGATGGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
 QY 349 CTACTTCAGAGCCAGCAGCTCCCGCCAGAGAGGCTGAGCCACACTGAGCCCGAGGAGC 408  
 Db 301 CTACTTCAGAGCCAGCAGCTCCCGCCAGAGAGGCTGAGCCACACTGAGCCCGAGGAGC 360  
 QY 409 AGGTTCTCTGGAGGAGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGT 468  
 Db 361 AGGTTCTCTGGAGGAGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGT 420  
 QY 469 CCCTTCTCCATGAATGGTACAGCGAGAACATG 501  
 Db 421 CCCTTCTCCATGAATGGTACAGCGAGAACATG 453

RESULT 47  
 CN483754  
 LOCUS  
 DEFINITION  
 hw36h12.y1 Human primary human ocular pericytes. Unamplified (hw)  
 Homo sapiens cDNA clone hw36h12 5', mRNA sequence.  
 CN483754  
 ACCESSION  
 CN483754.1 GI:46565258  
 VERSION  
 EST.  
 KEYWORDS  
 SOURCE  
 Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1. (bases 1 to 629)  
 AUTHORS  
 Tsai,J.F. and Wistow,G.  
 TITLE  
 Expressed sequence tag analysis of cultured primary human ocular  
 pericytes  
 JOURNAL  
 Unpublished (2004)  
 COMMENT  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 36 row: h column: 12  
 Seq primer: M13RPI reverse primer (ABI).  
 Location/Qualifiers  
 1. 629  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone\_lib="hw36h12"  
 /cell\_type="pericytes"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human primary human ocular pericytes."  
 Unamplified (hw)  
 /note="Organ: Eye; Vector: pSPORT1; RNA was extracted from  
 primary human pericytes in culture. A directionally cloned  
 cDNA library in the pSPORT1 vector (Invitrogen) was  
 constructed at Bioserve Biotechnology (Laurel MD)  
 essentially following the protocols of which are contained in the  
 Plasmid System full details of which are contained in the  
 manufacturer's instruction manual  
 (http://www.lifetech.com/). First strand synthesis was  
 carried out using a Not I primer-adaptor  
 [5'-pGACGTGTTCTAGATCGGAGCGCGCC(T)15-3']. cDNA was  
 cloned in Not I/Sal I sites. EST analysis was performed at  
 the NIH Intramural Sequencing Center (NISC)."

ORIGIN  
 Query Match 17.1%; Score 397; DB 7; Length 629;  
 Best Local Similarity 99.8%; Pred. No. 3.1e-201;  
 Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 985 AAGTTAAGAAAAAGAGCCCTAAACTTTTAAATAAATTTGATAAGACTATTAAAGCTGAAC 1044  
 Db 182 AAGTTAAGAAAAAGAGCCCTAAACTTTTAAATAAATTTGATAAGACTATTAAAGCTGAAC 241  
 QY 1045 TTGATGCTGCAGAAAAAATCCGTAAGAGGGGAAAAATTTGAGGAGCAGTCAATGCATTTA 1104  
 Db 242 TTGATGCTGCAGAAAAAATCCGTAAGAGGGGAAAAATTTGAGGAGCAGTCAATGCATTTA 301  
 QY 1105 AAGAACTAGTACGCAAAATACCTCAGAGTCCAGAGCAAGATATGGGAGCGCAGTGTG 1164  
 Db 302 AAGAACTAGTACGCAAAATACCTCAGAGTCCAGAGCAAGATATGGGAGCGCAGTGTG 361  
 QY 1165 AGGATGATTTGGCTGAGAGGAGGAGAAATGATAGGTGCTACGTGGAGCCATCGAGACCT 1224  
 Db 362 AGGATGATTTGGCTGAGAGGAGGAGAAATGATAGGTGCTACGTGGAGCCATCGAGACCT 421  
 QY 1225 ACCAAGAGGTGGCCAGGCTACCTGATGTCCTCGAGACCTGCTCAAGCTGAGTTTGAAGC 1284  
 Db 422 ACCAAGAGGTGGCCAGGCTACCTGATGTCCTCGAGACCTGCTCAAGCTGAGTTTGAAGC 481  
 QY 1285 GTCGCTCAGACAGCAACAATTTCTAGTCAATATGAGAGTTCCCTGCTACCTCGCAGAGA 1344  
 Db 482 GTCGCTCAGACAGCAACAATTTCTAGTCAATATGAGAGTTCCCTGCTACCTCGCAGAGA 541  
 QY 1345 GATTAGTTCAACTATTTCCCAATGATCTTCTTAAATAAATGACCTTGGCGTGGATACC 1404  
 Db 542 GATTAGTTCAACTATTTCCCAATGATCTTCTTAAATAAATGACCTTGGCGTGGATACC 601

```
QY 1405 TCTTGATAGGAGATATGACAAATGCAAA 1432
Db 602 TCTTGATAGGAGATATGACAAATGCAAA 629

RESULT 48
BQ340475/c
LOCUS BQ340475 497 bp mRNA linear EST 20-MAY-2002
DEFINITION OV2-NN2006-110501-627-c09 NN2006 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ340475
VERSION BQ340475.1 GI:21001375
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 497)
Nagai, M.A., da Silva, W. Jr., Zagor, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-NN2006-
110501-627-c09&t3=2001-05-11&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 496.
Location/Qualifiers
1..497
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN2006"
/notes="Organ: nervous normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORFESTES PCR (0.8. Letters Patent application
No.196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
```

```
QY 360 CCAGCATGTCGCGCAGACAGGCTGAGCCACACACTGAGCCGAGAGAGAGGTTCTCTGTG 419
Db 497 CCAGCATGTCGCGCAGACAGGCTGAGCCACACACTGAGCCGAGAGAGAGGTTCTCTGTG 438

QY 420 GAGCAGACCCAGCAATATCGAATGAGCAAAAGACAAATTCAGTCCCTTCTCCAT 479
Db 437 GAGCAGACCCAGCAATATCGAATGAGCAAAAGACAAATTCAGTCCCTTCTCCAT 378

QY 480 GAAATGGTACAGCAGACATGTTGAGGGAGAGACTTGCAACAGAGATGAGCCACA 539
|||||
```

```
Db 377 GAAATGGTACAGCAGACATGTTGAGGGAGAGAGACTTGCACAGAGATGAGCCACA 318
QY 540 GGAGAACCAACCAAGAGGATGATGAGTTCTTATGGCGACTGATGATGATGATGATGAT 599
Db 317 GGAGAACCAACCAAGAGGATGATGAGTTCTTATGGCGACTGATGATGATGATGATGAT 258
QY 600 GAGACCCCTGGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
Db 257 GAGACCCCTGGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 198
QY 660 ACAGTTTTCACAAAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGAT 719
Db 197 ACAGTTTTCACAAAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGAT 138
QY 720 GATTCAGTGAACCAAGCTAGTGAAGATGGAAGATGGAAGATGGAAGATGGAAGATGGAAG 779
Db 137 GATTCAGTGAACCAAGCTAGTGAAGATGGAAGATGGAAGATGGAAGATGGAAGATGGAAG 78
QY 780 TACCAAGTCTATGAGGAACCAAGCAGTATATGAACTCTAGAAAATGAAGGATAGAAATC 839
Db 77 TACCAAGTCTATGAGGAACCAAGCAGTATATGAACTCTAGAAAATGAAGGATAGAAATC 18
QY 840 ACAGAAGTAACTGCTCC 856
Db 17 ACAGAAGTAACTGCTCC 1

RESULT 49
BQ339559 503 bp mRNA linear EST 20-MAY-2002
LOCUS QV2-NN0045-101100-409-d06 NN0045 Homo sapiens cDNA, mRNA sequence.
DEFINITION QV2-NN0045-101100-409-d06 NN0045 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ339559
VERSION BQ339559.1 GI:20999368
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 503)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zagor, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-NN0045-
101100-409-d06&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 4.
Location/Qualifiers
1..503
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN0045"
/notes="Organ: nervous normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
```



This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&ts=QV2-NN2003-230401-624-a03&t3=2001-04-23&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 525.

## FEATURES

Location/Qualifiers  
 1..553  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="NN2003"  
 /note="Organ: nervous normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 16.7%; Score 389; DB 4; Length 553;  
 Best Local Similarity 99.6%; Pred. No. 66-197;  
 Matches 489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 332 AGGACTTAAAGAGAGATCTACTTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 391  
 DB 17 AGGACTTAAAGAGAGATCTACTTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 76

QY 392 CACTGAGCCGAGGAGCAGGTTCTCTGTGGAGGAGCAACCCAGAGATATCGAAGTGAAGC 451  
 DB 77 CACTGAGCCGAGGAGCAGGTTCTCTGTGGAGGAGCAACCCAGAGATATCGAAGTGAAGC 136

QY 452 AAAAGAACAAATTCAGTCCCTCTCCATGATGTGTACAGCGAGACATGTTGAGGGAGA 511  
 DB 137 AAAAGAACAAATTCAGTCCCTCTCCATGATGTGTACAGCGAGACATGTTGAGGGAGA 196

QY 512 AGACTTGCACAGAGATGAGCCACAGGAGAACCCACACAGAGAGATGATGATTTCT 571  
 DB 197 AGACTTGCACAGAGATGAGCCACAGGAGAACCCACACAGAGAGATGATGATTTCT 256

QY 572 TATGGCAGTGTATGATGATGATGATTTGAGACCTCGAAGTATCTCATGAAGA 631  
 DB 257 TATGGCAGTGTATGATGATGATGATTTGAGACCTCGAAGTATCTCATGAAGA 316

QY 632 AACCGACATAGTTTACCAGTGAAGAGACAGTTCACAGAGCTGTAATCAGGATATGGA 691  
 DB 317 AACCGACATAGTTTACCAGTGAAGAGACAGTTCACAGAGCTGTAATCAGGATATGGA 376

QY 692 AGAGATGATCTGTGACGAGAAATCCAGATTCAGTGAACCGAGTGTAGAGATGAAG 751  
 DB 377 AGAGATGATCTGTGACGAGAAATCCAGATTCAGTGAACCGAGTGTAGAGATGAAG 436

QY 752 ATTGCACATGATACAGATGATGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGA 811  
 DB 437 ATTGCACATGATACAGATGATGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGA 496

QY 812 ACCTCTAGAAA 822  
 DB 497 ACCTCTAGAAA 507

## RESULT 52

BE084010/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 BE084010  
 VERSION  
 BE084010.1  
 GI:8474363  
 EST.  
 Homo sapiens (human)  
 SOURCE  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

BE084010 409 bp mRNA linear EST 12-JUN-2000  
 PMO-BT0651-150300-002-905 BT0651 Homo sapiens cDNA, mRNA sequence.

## REFERENCE

## AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 409)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## MEDLINE

20202663

## PUBMED

10737800

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil.  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=PMO-BT0651-150300-002-g05&t3=2000-03-15&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 18

High quality sequence stop: 409.

## FEATURES

Location/Qualifiers

1..409

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="BT0651"

/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 16.6%; Score 385; DB 2; Length 409;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-195;  
 Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 ATCTACTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGA 406  
 DB 402 ATCTACTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGA 343

QY 407 GCAGGTTCTGTGGAGGAGCAACCCAGAGATATCGAAGATGAAGCAAAAGCAAAATTC 466  
 DB 342 GCAGGTTCTGTGGAGGAGCAACCCAGAGATATCGAAGATGAAGCAAAAGCAAAATTC 283

QY 467 GTCCCTTCTCCATGAATGTGTACAGCGAGAACCATGTTGAGGAGAGAGTTCGACACAGA 526  
 DB 282 GTCCCTTCTCCATGAATGTGTACAGCGAGAACCATGTTGAGGAGAGAGTTCGACACAGA 223

QY 527 AGATGAGCCAGGAGCAACCAAGAGGATGATGATTTCTTATGGCGACTGATGT 586  
 DB 222 AGATGAGCCAGGAGCAACCAAGAGGATGATGATTTCTTATGGCGACTGATGT 163

QY 587 AGATGATAGATTGAGACCCCTGGAACTCTGAGTATCTCATGAAGAAACCGAGCATAGTTA 646  
 DB 162 AGATGATAGATTGAGACCCCTGGAACTCTGAGTATCTCATGAAGAAACCGAGCATAGTTA 103

QY 647 CCACGTGGAAGAGACAGTTTTCACAAAGTCTGTAATCAGGATATGGAAGAGATGATCTCTGA 706  
 DB 102 CCACGTGGAAGAGACAGTTTTCACAAAGTCTGTAATCAGGATATGGAAGAGATGATCTCTGA 43

QY 707 GCAGGAAATTCAGATTCAGTGA 731

Db	318	GGTTTATGGTGATTCATTTGCTGGCGCTCTGGACATCTCTAGCTGTCGCTTTGGTTTGATC	377
Qy	241	TTGTTGACTATGAGGAAGTTCTAGGAAAATAGGAATCTATGATGCTGATGCTGATGGAG	300
Db	378	TTGTTGACTATGAGGAAGTCTTAGGAAAATAGGAATCTATGATGCTGATGCTGATGGAG	437
Qy	301	ATTTTCATGTGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360
Db	438	ATTTTATGATGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	497
Qy	361	CAGCAGTCCCGCAGGAAGGCTGAGCCACACACTGAGCCCGAGGAGAGAGTTCCTGTGG	420
Db	498	CAGCAGTCCCGCAGGAAGGCTGAGCCACACACTGAGCCCGAGGAGAGAGTTCCTGTGG	557
Qy	421	AGCAGAACCCAGATATCGAGATGAAGCAAAAGACAAATTCAGTCCCTTCTCCATG	480
Db	558	AGCAGAACCCAGATATCGAGATGAAGCAAAAGACAAATTCAGTCCCTTCTCCATG	617
Qy	481	AAATGGTACACGAGAACATGTTGAGGAGAGACTTGCAACAAGAGATGGACCC	536
Db	618	AAATGGTACACGAGAACATGTTGAGGAGAGACTTGCAACAAGAGATGGACCC	673
RESULT	54		
CN368831			
LOCUS		468 bp	linear
DEFINITION		1700600846567 GRN_PRENU Homo sapiens cDNA 5', mRNA sequence.	EST 16-MAY-2004
ACCESSION		CN368831	
VERSION		CN368831.1	GI:47368765
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J. and Stanton, L. W.	
TITLE		Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation	
JOURNAL		Nat. Biotechnol. 22 (6), 707-716 (2004)	
COMMENT		Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 468 Std Error: 0.00.	
FEATURES		Location/Qualifiers	
source		1..468	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/tissue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"	
		/clone_lib="GRN PRENU"	
		/note="Oligo dT primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic acid and mitogens."	
ORIGIN			
Query Match		16.4%; Score 382; DB 7; Length 468;	
Best Local Similarity		100.0%; Pred. No. 3.4e-193;	
Matches		382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	251	TCAGGAAGTCTTAGGAAACCTAGGAATCTATGATGCTGATGGATGGAGATTTTGATGT	310
Db	7	TCAGGAAGTCTTAGGAAACCTAGGAATCTATGATGCTGATGGATGGAGATTTTGATGT	66
Qy	311	CGATGATCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGCCAGAGTCCC	370

Db 67 GGATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCC 126  
 Qy 371 GCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCTGTGGAGGCGAGAAC 430  
 Db 127 GCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCTGTGGAGGCGAGAAC 186  
 Qy 431 CCAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCCTCATGAAATGCTACA 490  
 Db 187 CCAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCCTCATGAAATGCTACA 246  
 Qy 491 CCAGAGACATGTTGAGGAGAGAGACTTGCACAGAGAGATGAGCCACAGAGAGAACCA 550  
 Db 247 CCAGAGACATGTTGAGGAGAGAGACTTGCACAGAGAGATGAGCCACAGAGAGAACCA 306  
 Qy 551 ACAGAGGATGATGAGTTTCTTATGCGAGCTGATGTAGATGATAGATTGAGACCCCTGGA 610  
 Db 307 ACAAGAGGATGATGAGTTTCTTATGCGAGCTGATGTAGATGATAGATTGAGACCCCTGGA 366  
 Qy 611 ACCTGAAGTATCTCATGAAGAA 632  
 Db 367 ACCTGAAGTATCTCATGAAGAA 388

RESULT 55  
 LOCUS CA311352 720 bp mRNA linear EST 04-NOV-2002  
 DEFINITION UI-CF-FNO-aff-d-19-0-UI.81 UI-CF-FNO Homo sapiens cDNA clone  
 ACCESSION CA311352  
 VERSION CA311352.1 GI:24529450  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 720)  
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 9704477  
 PUBMED 889548  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com) or from Open Biosystems  
 (www.openbiosystems.com).  
 Seq primer: M13 FORWARD  
 POLYA=Yes.  
 Location/Qualifiers  
 1. 720  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue="UI-CF-FNO-aff-d-19-0-UI"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-CF-FNO"  
 /notes="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a  
 modified polylinker; Site: 1: EcoR I; Site 2: Not I;  
 UI-CF-FNO is a subtracted cDNA library derived from two  
 normalized Human lung epithelial cell libraries (EN1 and  
 DU1) The library was subtracted according to according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:  
 bent-soares@uiowa.edu  
 TAG\_L18SUB=lung Epithelial Cells Tissue nos 359-368  
 TAG\_L18=UI-CF-FNO  
 TAG\_SEQ=GGCTGTAGGC"

## FEATURES

source

## ORIGIN

Query Match 16.3%; Score 379; DB 6; Length 720;  
 Best Local Similarity 99.8%; Pred. No. 1.5e-191;  
 Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1895 CACGCTGTGGCAGCAAGGAAGAAGAAATGAAATSCCTGCAAGAGAGCTCTCTAAAACCTG 1954  
 Db 51 CACGCTGTGGCAGCAAGGAAGAAGAAATGAAATSCCTGCAAGAGAGCTCTCTAAAACCTG 110  
 Qy 1955 TACCTTACTAGAAAGTTCCCGAGACACACAGGATGCGAGAGAGAGACAGATCAATATTC 2014  
 Db 111 TACCTTACTAGAAAGTTCCCGAGACACACAGGATGCGAGAGAGAGACAGATCAATATTC 170  
 Qy 2015 CATCATGACCCCGGAGACTCACGTGTGCGCCACACAGGCCCCCAAACTGCAGGCTCG 2074  
 Db 171 CATCATGACCCCGGAGACTCACGTGTGCGCCACACAGGCCCCCAAACTGCAGGCTCG 230  
 Qy 2075 AATGACACTGGGCTTGGTGATTCCTCAAGGAAGGCTGCAAGATTGTCCTCAACGAGAC 2134  
 Db 231 AATGACACTGGGCTTGGTGATTCCTCAAGGAAGGCTGCAAGATTGTCCTCAACGAGAC 290  
 Qy 2135 CAGGACCTGGGAGGAAAGGAGGAGTGTCTCATCTTTGATGACTCTTTGAGCAGGAGTATG 2194  
 Db 291 CAGGACCTGGGAGGAAAGGAGGAGTGTCTCATCTTTGATGACTCTTTGAGCAGGAGTATG 350  
 Qy 2195 GCAGATGCTCTATCTTTCCGCTGATATTCATCGTGGATGTGGCATCCGGAACCTGAC 2254  
 Db 351 GCAGATGCTCTATCTTTCCGCTGATATTCATCGTGGATGTGGCATCCGGAACCTGAC 410  
 Qy 2255 ACCACACAGAGAGCGAGCCTTCCAGCAATTTAGCATGAATTCATCAAGCTTTGGGAAC 2314  
 Db 411 ACCACACAGAGAGCGAGCCTTCCAGCAATTTAGCATGAATTCATCAAGCTTTGGGAAC 470  
 Qy 2315 TCTGGAGAGA 2324  
 Db 471 TCTGGAGAGA 480

## RESULT 56

LOCUS CN368826 626 bp mRNA linear EST 16-MAY-2004  
 DEFINITION 17000423444515 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
 ACCESSION CN368826  
 VERSION CN368826.1 GI:47368760  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 626)  
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
 Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
 Lebkowski, J. and Stanton, L.W.  
 REFERENCE Transcriptionome characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 TITLE Contact: Brandenberger R  
 JOURNAL Regenerative Medicine  
 COMMENT Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@geron.com  
 Insert Length: 626 Std Error: 0.00.  
 LOCATION/Qualifiers  
 1. 626  
 /organism="Homo sapiens"



/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/notes="oligo dt primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

ORIGIN

Query Match 15.9%; Score 370; DB 7; Length 626;  
Best Local Similarity 99.8%; Pred. No. 1e-186;  
Matches 420; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGACAGCGGCAACAGCAGCAGCG 60  
Db 206 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGACAGCGGCAACAGCAGCAGCG 265  
Qy 61 GCTCCGCGAGGTAGCAGCAGTGGCGCAGCAGCAGCGGCGGCGGCGGCGGCGGCGG 120  
Db 266 GCTCCGCGAGGTAGCAGCAGTGGCGCAGCAGCAGCGGCGGCGGCGGCGGCGGCGG 325  
Qy 121 AGCATGGAGGACACAAGATGGGAGGAAAGCGGACTCTCGGGAACCTTCATCTTCACGT 180  
Db 326 AGCATGGAGGACACAAGATGGGAGGAAAGCGGACTCTCAGGAACTTCATCTTCACGT 385  
Qy 181 GGTATTAGTGAATGTCATGCTGGGCGTCTGGACATCTGTAGCTGTCTGTGTTGATC 240  
Db 386 GGTATTAGTGAATGTCATGCTGGGCGTCTGGACATCTGTAGCTGTCTGTGTTGATC 445  
Qy 241 TTGTTGACTATGAGGAGTCTAGGAACTAGGAATCTATGATGCTGATGCTGATGGAG 300  
Db 446 TTGTTGACTATGAGGAGTCTAGGAACTAGGAATCTATGATGCTGATGCTGATGGAG 505  
Qy 301 ATTTTGTGATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Db 506 ATTTTGTGATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 565  
Qy 361 CAGCAGTCCCGCAGAGGCTGAGCAGACACACTGAGCCGCGAGGAGGAGTCTCTGTTGG 420  
Db 566 CAGCAGTCCCGCAGAGGCTGAGCAGACACACTGAGCCGCGAGGAGGAGTCTCTGTTGG 625  
Qy 421 A 421  
Db 626 A 626

RESULT 57  
BE876059  
LOCUS  
DEFINITION  
mRNA sequence.  
601485742F1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3887962 5',  
1048 bp mRNA linear EST 20-OCT-2000  
ACCESSION  
BE876059  
VERSION  
BE876059.1 GI:10324835  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1048)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: DCTD/DRP/Gazdar  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLAM9667 row: d column: 11  
High quality sequence stop: 670.  
Location/Qualifiers  
1. 1048  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3887962"  
/tissue\_type="large cell carcinoma, undifferentiated"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 69"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.1 Kb. Library constructed by Life  
Technologies."

ORIGIN

Query Match 14.2%; Score 329; DB 2; Length 1048;  
Best Local Similarity 99.5%; Pred. No. 1.3e-164;  
Matches 429; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGACAGCGGCAACAGCAGCAGCG 60  
Db 169 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGACAGCGGCAACAGCAGCAGCG 228  
Qy 61 GCTCCGCGAGGTAGCAGCAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
Db 229 GCTCCGCGAGGTAGCAGCAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 288  
Qy 121 AGCATGGAGGACACAAGATGGGAGGAAAGCGGACTCTCGGGAACCTTCATCTTCACGT 180  
Db 289 AGCATGGAGGACACAAGATGGGAGGAAAGCGGACTCTCAGGAACTTCATCTTCACGT 348  
Qy 181 GGTATTAGTGAATGTCATGCTGGGCGTCTGGACATCTGTAGCTGTCTGTGTTGATC 240  
Db 349 GGTATTAGTGAATGTCATGCTGGGCGTCTGGACATCTGTAGCTGTCTGTGTTGATC 408  
Qy 241 TTGTTGACTATGAGGAGTCTAGGAACTAGGAATCTATGATGCTGATGCTGATGGAG 300  
Db 409 TTGTTGACTATGAGGAGTCTAGGAACTAGGAATCTATGATGCTGATGCTGATGGAG 468  
Qy 301 ATTTTGTGATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Db 469 ATTTTGTGATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 528  
Qy 361 CAGCAGTCCCGCAGAGGCTGAGCAGACACACTGAGCCGCGAGGAGGAGTCTCTGTTGG 420  
Db 529 CAGCAGTCCCGCAGAGGCTGAGCAGACACACTGAGCCGCGAGGAGGAGTCTCTGTTGG 588  
Qy 421 AGGAGAGACCC 431  
Db 589 AGGAGAGACCC 599

RESULT 58

AI267311/c  
LOCUS  
DEFINITION  
aq63e03.x1 Stanley Frontal SN pool 2 Homo sapiens cDNA clone  
IMAGE:2035612 similar to SW:ASPH HUMAN Q12797 ASPARYL/ASPARAGINYL  
BETA-HYDROXYLASE ; mRNA sequence.  
AI267311  
ACCESSION  
AI267311.1 GI:3886478  
VERSION  
EST.  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 574)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,  
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST Project  
TITLE

JOURNAL  
COMMENT

Unpublished (1997)  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewartson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -400P from Gibco  
High quality sequence stop: 439.

FEATURES  
source

Location/Qualifiers  
1..574  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2035612"  
/tissue types="frontal lobe (see description)"  
/clone host="Stanley Frontal SN pool 2"  
/note="Organ: brain; Vector: pOR2.1-TOPO (Invitrogen);  
before use) was reverse transcribed using a modified  
oligo-dr primer containing RsaI and HindIII sites.  
Double-stranded cDNA was digested with RsaI, resulting in  
blunt ended cDNA of an average 0.1-2 kb in length.  
Digested cDNA was split into two sets, one used as is as  
the driver, the other set was split in half again and each  
half linked to a different adaptor  
(5'-TCGAGCGCGCGCGCGAGGT-3' or 5'-  
AGGCGGTGTCGCGAGCGGT-3'), to be used as tester.  
Subtraction was performed using the Clontech PCR Select  
cDNA subtraction kit. Pool of two schizoprenics, male age  
44 and female age 56 (S-116, S-118) subtracted by pool of  
two mentally normal male individuals ages 41 and 53  
(S-124, S-141). Tissues were obtained from the Stanley  
Neuropathology Consortium (www.stanleylab.org). Library  
constructed and subtracted by Dr. Nancy Johnston [(410)  
614-3918, nlj@welchlink.welch.jhu.edu]."

## ORIGIN

Query Match 14.1%; Score 328; DB 1; Length 574;  
Best Local Similarity 100.0%; Pred. No. 4.1e-164;  
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCACGTGTTATGGTGATTCGATGCGGCGTCTGGACATCTGTA 221  
DB 345 GGAATTCATCTTCACGTGTTATGGTGATTCGATGCGGCGTCTGGACATCTGTA 286  
QY 222 GCTGCTGTTGGTTCATCTGTCATGATGAGGAGTCTAGAAAACCTAGGAATCTAT 281  
DB 285 GCTGTGTTGGTTCATCTGTCATGATGAGGAGTCTAGAAAACCTAGGAATCTAT 226  
QY 282 GATGCTGATGATGAGGAGTTCATGATGATGATGATGATGATGATGATGATGATGAT 341  
DB 225 GATGCTGATGATGAGGAGTTCATGATGATGATGATGATGATGATGATGATGATGAT 166  
QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGGCC 401  
DB 165 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGGCC 106  
QY 402 GAGGAGCAGGTTCTCTGGAGCAGACCCAGAGATATCGAATGAGCAAGCAACAA 461  
DB 105 GAGGAGCAGGTTCTCTGGAGCAGACCCAGAGATATCGAATGAGCAAGCAACAA 46  
QY 462 ATTCAGTCCCTTCTCCATGAATGTAC 489  
DB 45 ATTCAGTCCCTTCTCCATGAATGTAC 18

RESULT 59  
CN368818  
LOCUS

CN368818 393 bp mRNA linear EST 16-MAY-2004

## DEFINITION

17000600014160 GRN\_PRENU Homo sapiens cDNA 5', mRNA sequence.  
CN368818  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)

## REFERENCE

1 (bases 1 to 393)  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
Lebkowski, J. and Stanton, L.W.  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation

## AUTHORS

Nat. Biotechnol. 22 (6), 707-716 (2004)  
Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 393 Std Error: 0.00.

## JOURNAL

## COMMENT

## FEATURES

## source

## Location/Qualifiers

## 1..393

## /organism="Homo sapiens"

## /mol\_type="mRNA"

## /db\_xref="taxon:9606"

/tissue\_type="embryonic stem cell, retinoic acid and  
mitogen-treated HES cell line H7"

## /clone\_lib="GRN\_PRENU"

/notes="Oligo dT primed, full-length enriched cDNA library  
from HES cell line H7 (p29) maintained in feeder-free  
conditions. Embryoid bodies were generated in the presence  
of all-trans retinoic acid and mitogens."

## ORIGIN

Query Match 13.9%; Score 324; DB 7; Length 393;  
Best Local Similarity 100.0%; Pred. No. 5.4e-162;  
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCACGTGTTATGGTGATTCGATGCGGCGTCTGGACATCTGTA 221  
DB 70 GGAATTCATCTTCACGTGTTATGGTGATTCGATGCGGCGTCTGGACATCTGTA 129  
QY 222 GCTGCTGTTGGTTCATCTGTCATGATGAGGAGTTCAGAAAACCTAGGAATCTAT 281  
DB 130 GCTGCTGTTGGTTCATCTGTCATGATGAGGAGTTCAGAAAACCTAGGAATCTAT 189  
QY 282 GATGCTGATGATGAGGAGTTCATGATGATGATGATGATGATGATGATGATGATGAT 341  
DB 190 GATGCTGATGATGAGGAGTTCATGATGATGATGATGATGATGATGATGATGATGAT 249  
QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGGCC 401  
DB 250 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGGCC 309  
QY 402 GAGGAGCAGGTTCTCTGGAGCAGACCCAGAGATATCGAATGAGCAAGCAACAA 461  
DB 310 GAGGAGCAGGTTCTCTGGAGCAGACCCAGAGATATCGAATGAGCAAGCAACAA 369  
QY 462 ATTCAGTCCCTTCTCCATGAATG 485  
DB 370 ATTCAGTCCCTTCTCCATGAATG 393

## RESULT 60

## CN481268

## LOCUS

## DEFINITION

## Homo sapiens cDNA clone hw05d04 5', mRNA sequence.

## ACCESSION

## VERSION

CN481268 580 bp mRNA linear EST 26-APR-2004  
hw05d04.y1 Human primary human ocular pericytes. Unamplified (hw)  
Homo sapiens cDNA clone hw05d04 5', mRNA sequence.  
CN481268  
CN481268.1 GI:46562772

KEYWORDS	EST.	Db	515	CAGCAGTCCCGCC	527
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 580)				
AUTHORS	Tsai,J.Y. and Wistow,G.				
TITLE	Expressed sequence tag analysis of cultured primary human ocular pericytes				
JOURNAL	Unpublished (2004)				
COMMENT	Contact: Wistow G Section on Molecular Structure and Function National Eye Institute 6/331, NIH, Bethesda, MD 20892-2740, USA Tel: 301 402 3452 Fax: 301 496 0078 Email: graeme@helix.nih.gov Plate: 05 row: d column: 04 Seq primer: M13Rpl reverse primer (ABI). Location/Qualifiers				
FEATURES	1..580 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="hw05d04" /cell_type="pericytes" /dev_stage="Adult" /lab_host="EMDH10B" /clone_lib="Human primary human ocular pericytes. Unamplified (hw)" /note="Organ: Eye; Vector: pSport1; RNA was extracted from primary human pericytes in culture. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGATAGTTCTAGATCGGAGCGGCCCT(T)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC)."				
source					
ORIGIN					
Query Match	13.9%; Score 322; DB 7; Length 580;				
Best Local Similarity	99.7%; Pred. No. 6.9e-161;				
Matches	372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
Qy	1 CGGACCGTGCATGGCCCAAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG 60				
Db	155 CGGACCGTGCATGGCCCAAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG 214				
Qy	61 GTTCCGGCAGCGGTAGCAGTGGGGCAGCAGCGCCCGGGCCCGGAGAGACAA 120				
Db	215 GTTCCGGCAGCGGTAGCAGTGGGGCAGCAGCGCCCGGGCCCGGAGAGACAA 274				
Qy	121 AGCATGAGGACACAAAGATGGAGAAAGCGGACTCTCGGGAACCTTCATTTTCAGGT 180				
Db	275 AGCATGAGGACACAAAGATGGAGAAAGCGGACTCTCGGGAACCTTCATTTTCAGGT 334				
Qy	181 GGTTTATGGTGAATTCGTTGGGCGTCTGGACATCTGTAGCTGTGTTGGTTGATC 240				
Db	335 GGTTTATGGTGAATTCGTTGGGCGTCTGGACATCTGTAGCTGTGTTGGTTGATC 394				
Qy	241 TTGTTCACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGTGTTGGTGGAG 300				
Db	395 TTGTTCACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGTGTTGGTGGAG 454				
Qy	301 ATTTTGTATGGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC 360				
Db	455 ATTTTGTATGGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC 514				
Qy	361 CAGCAGTCCCGCC 373				
DEFINITION	AL694159 527 bp mRNA linear EST 04-SEP-2003				
LOCUS	DKFZp313D1441.r1 313 (synonym: hlcc2) Homo sapiens cDNA clone				
ACCESSION	AL694159				
VERSION	AL694159.1 GI:19617857				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 527)				
AUTHORS	Koehler,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.				
TITLE	EST (Koehler,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: MIPS				
FEATURES	1..527 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DKFZp313D1441" /dev_stage="adult" /lab_host="DH10B" /clone_lib="313 (synonym: hlcc2)" /note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"				
source					
ORIGIN					
Query Match	13.8%; Score 321; DB 1; Length 527;				
Best Local Similarity	100.0%; Pred. No. 2.3e-160;				
Matches	321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	626 TGAAGAAACCGAGCATAGTTACCACTGGGAGAGACAGTTTCCACAGACTGTAATCAGGA 685				
Db	1 TGAAGAAACCGAGCATAGTTACCACTGGGAGAGACAGTTTCCACAGACTGTAATCAGGA 60				
Qy	686 TATGGAAGAGATGATGTCGACGAGGAAATCCAGATTCAGATGCAACAGTAGTAGAAGA 745				
Db	61 TATGGAAGAGATGATGTCGACGAGGAAATCCAGATTCAGATGCAACAGTAGTAGAAGA 120				
Qy	746 TGAAGATGTCACCATGATACAGATGATGTAACATACCAAGTCTATGAGGAACAGCAGT 805				
Db	121 TGAAGATGTCACCATGATACAGATGATGTAACATACCAAGTCTATGAGGAACAGCAGT 180				
Qy	806 ATATGAACCTCTAGAAATGAGGAGTAGAATCACAAGAGTAAGTCTCTCCCTCAGGA 865				
Db	181 ATATGAACCTCTAGAAATGAGGAGTAGAATCACAAGAGTAAGTCTCTCCCTCAGGA 240				
Qy	866 TAATCTCTAGAGATTTCACAGGTAATTTGTAGAAGAGTAAGCAATTTTCTGTGGAAGA 925				
Db	241 TAATCTCTAGAGATTTCACAGGTAATTTGTAGAAGAGTAAGCAATTTTCTGTGGAAGA 946				
Qy	926 ACAGCAGGAAGTACCACCAGA 946				
Db	301 ACAGCAGGAAGTACCACCAGA 321				

```

RESULT 62
AY405268      2023 bp      DNA      linear      GSS 16-DSB-2003
LOCUS      Pan troglodytes ASPH gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY405268
VERSION      AY405268.1      GI:39761242
KEYWORDS      GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM      Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE      1 (bases 1 to 2023)
AUTHORS      Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
              Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
              Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
              Adams, M.D. and Cargill, M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE      2 (bases 1 to 2023)
AUTHORS      Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
              Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
              Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
              Adams, M.D. and Cargill, M.
TITLE      Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT      This sequence was made by sequencing genomic exons and ordering
              them based on alignment
FEATURES      Location/Qualifiers
source      1..2023
              /organism="Pan troglodytes"
              /mol_type="genomic DNA"
              /db_xref="taxon:9598"
              <1..2023
              /gene="ASPH"
              /locus_tag="HCM2167"
genotype
ORIGIN
Query Match      13.6%;      Score 316;      DB 9;      Length 2023;
Best Local Similarity      99.7%;      Pred. No. 1.5e-157;
Matches 366;      Conservative      0;      Mismatches      1;      Indels      0;      Gaps      0;

QY      998      GAAGCCTAAACTTTTAAATAATTTGATAGACTATTAAAGCTGAACCTGATGCTGCAGAG      1057
DB      884      GAAGCCTAAACTTTTAAATAATTTGATAGACTATTAAAGCTGAACCTGATGCTGCAGAG      943

QY      1058      AAAACTCGTAAAGGGGAAAAATTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGC      1117
DB      944      AAAACTCGTAAAGGGGAAAAATTGAGGAAGCAGTGAATGCGTTTAAAGAACTAGTAGC      1003

QY      1118      CAATATACCTCAGAGTCCACGACGAGATATGGGAAGCGCAGTGTGAGGATGATTTGCG      1177
DB      1004      CAATATACCTCAGAGTCCACGACGAGATATGGGAAGCGCAGTGTGAGGATGATTTGCG      1063

QY      1178      TGAGAAGGAGAGAGTAAGTACGTGCTACGTGGAGCCATCGAGACCTACGAGAGGTGCG      1237
DB      1064      TGAGAAGGAGAGAGTAAGTACGTGCTACGTGGAGCCATCGAGACCTACGAGAGGTGCG      1123

QY      1238      CAGCTACCTGATGTCCTCGACAGCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAG      1297
DB      1124      CAGCTACCTGATGTCCTCGACAGCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAG      1183

QY      1298      GCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCTCGACAGATTTAGTTCAACT      1357
DB      1184      GCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCTCGACAGATTTAGTTCAACT      1243

QY      1358      ATTTCCC      1364
|||||

```

```

Db      1244      ATTTCCC      1250

RESULT 63
BX111251
LOCUS      BX111251
DEFINITION      IMAGE:1876994, mRNA sequence.
ACCESSION      BX111251
VERSION      BX111251.1      GI:27878432
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 484)
AUTHORS      Ebert, D., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
              Radelof, U., Schneider, D. and Korn, B.
TITLE      Human UnigeneSet - RZPD3
JOURNAL      Unpublished (2003)
COMMENT      Contact: Ina Rolfs
              RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
              Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
              RZPD; IMAGP998F034597.
              RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
              Human UnigeneSet - RZPD3 (RZPDLIB No.972)
              http://www.rzpd.de/CloneCards/cgi-
              bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
              RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
              Heubnerweg 6, D-14059 Berlin, Germany
              Tel: +49 30 32639 101
              Fax: +49 30 32639 111
              www.rzpd.de
              This clone is available royalty-free from RZPD;
              contact RZPD (clone@rzpd.de) for further information. Seq primer:
              M13r, Primer sequence: TTTCACACAGAAACAGCTATGAC.
FEATURES      Location/Qualifiers
source      1..484
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGP998F034597 ; IMAGE:1876994"
              /tissue_type="pooled human melanocyte, fetal heart, and
              pregnant uterus"
              /lab_host="DH10B"
              /clone_lib="Soares_NHMPu_S1"
              /notes="Organ: mixed (see below); Vector: pT7T3D-Pac
              (Pharmacia) with a modified polylinker; Site 1: Not I;
              Site 2: Eco RI; Equal amounts of plasmid DNA from three
              normalized libraries (melanocyte 2NBHM, pregnant uterus
              NBHPU, and fetal heart NBHH19W) were mixed, and ss circles
              were made in vitro. Following HAP purification, this DNA
              was used as tracer in a subtractive hybridization
              reaction. The driver was PCR-amplified cDNAs from pools of
              5,000 clones made from the same 3 libraries. The pools
              consisted of I.M.A.G.E. clones 260232-265223,
              340488-345479, and 484488-489479."
ORIGIN
Query Match      13.5%;      Score 313;      DB 5;      Length 484;
Best Local Similarity      100.0%;      Pred. No. 4.7e-156;
Matches 313;      Conservative      0;      Mismatches      0;      Indels      0;      Gaps      0;

QY      848      AACTGCTCCCCCTGAGGATAATCTGTAGAGATTTCACAGGTAAATGTGTAAGAAAGTAAG      907
DB      1      AACTGCTCCCCCTGAGGATAATCTGTAGAGATTTCACAGGTAAATGTGTAAGAAAGTAAG      60

QY      908      CATTTTTCCTGTGTGGAAGAACACGAGAGTACACCAAGAACAAATAGAAAACAGATGA      967
DB      61      CATTTTTCCTGTGTGGAAGAACACGAGAGTACACCAAGAACAAATAGAAAACAGATGA      120

QY      968      TCCAGACAAAAGCAAAAGTTAGAAAAGAGAGCCCTAAACTTTTAAATTTGATTA      1027
|||||

```

```

Db      121  TCAGAGCAAAAGCAAAAGTTAAGAAAAAGAGCCCTAAACCTTTTAAATAAATTTGATA 180
Qy      1028 GACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAAGCTCCGTAAAGGGGAAAAATTCAGGA 1087
Db      181  GACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAAGCTCCGTAAAGGGGAAAAATTCAGGA 240
Qy      1088 AGCAGTGAATGCTATTAAAGAACTAGTAGCAGCAATACCCCTCAGAGTCCAGGCAAGATA 1147
Db      241  AGCAGTGAATGCTATTAAAGAACTAGTAGCAGCAATACCCCTCAGAGTCCAGGCAAGATA 300
Qy      1148 TGGGAAGCGCGAG 1160
Db      301  TGGGAAGCGCGAG 313

RESULT 64
CF136513
LOCUS
DEFINITION
UI-HF-BNO-amb-d-05-0-UI.r1 NIH MGC_50 Homo sapiens cDNA clone
IMAGE:3092001 5', mRNA sequence.
ACCESSION
CF136513
VERSION
CF136513.1 GI:33251957
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 503)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Louis Staudt
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.
FEATURES
Location/Qualifiers
1..503
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3092001"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="MGC85"
/lab_host="DH10B (LTI)"
/clone_lib="NIH MGC 50"
/note="Vector: pYX3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
ORIGIN
Query Match 13.4%; Score 312; DB 7; Length 503;
Best Local Similarity 99.5%; Pred. No. 1.6e-155;
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      907 GCATTTTCTCTGGGAAGAACAGCAGGAAGTACCACGAGAAACAATAGAAAAACAGATG 966
Db      59 GCATTTTCTCTGGGAAGAACAGCAGGAAGTACCACGAGAAACAATAGAAAAACAGATG 118

```

```

Qy      967 ATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAAACCTTTTAAATAAATTTGATA 1026
Db      119  ATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAAACCTTTTAAATAAATTTGATA 178
Qy      1027 AGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAAGCTCCGTAAAGGGGAAAAATTCAGG 1086
Db      179  AGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAAGCTCCGTAAAGGGGAAAAATTCAGG 238
Qy      1087 AAGCAGTGAATGCTATTAAAGAACTAGTAGCAGCAATACCCCTCAGAGTCCAGGCAAGAT 1146
Db      239  AAGCAGTGAATGCTATTAAAGAACTAGTAGCAGCAATACCCCTCAGAGTCCAGGCAAGAT 298
Qy      1147 ATGGGAAGCGCGAGTGTGAGGATGATTTGGCTGAGAAGAGGAGCAAGTAATAGGTCCTAC 1206
Db      299  ATGGGAAGCGCGAGTGTGAGGATGATTTGGCTGAGAAGAGGAGCAAGTAATAGGTCCTAC 358
Qy      1207 GTGGAGCCATCGAGACCTACCAAGAGTGGCAGGCTTACTGATGTCCTGCAGACCTGC 1266
Db      359  GTGGAGCCATCGAGACCTACCAAGAGTGGCAGGCTTACTGATGTCCTGCAGACCTGC 418
Qy      1267 TGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGTCTCATATGA 1320
Db      419  TGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGTCTCATATGA 472

RESULT 65
BG170148
LOCUS
DEFINITION
602321860F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4424894 5',
mRNA sequence.
ACCESSION
BG170148
VERSION
BG170148.1 GI:12676936
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 682)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: sgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10169 row: h column: 15
High quality sequence stop: 682.
FEATURES
Location/Qualifiers
1..682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4424894"
/tissue_type="hypermphroma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 89"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.3 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 13.4%; Score 311; DB 4; Length 682;
Best Local Similarity 100.0%; Pred. No. 5.9e-155;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor 15'-pGACTTCGTAGTTCGGAGCGGCCCGCC(T)15-3'. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 13.2%; Score 306; DB 7; Length 560;  
Best Local Similarity 99.7%; Pred. No. 2.8e-152; Indels 0; Gaps 0;  
Matches 356; Conservative 0; Mismatches 1;

QY 1 CGGACCGTCAATGGCCAGCGTAAGATGCCAAGCAGCGCGCAACAGCAGCAGCG 60  
Db 204 CGGACCGTCAATGGCCAGCGTAAGATGCCAAGCAGCGCGCAACAGCAGCAGCG 263  
QY 61 GTCCGGCAGCGGTAGCAGTGGCGGAGCAGCAGCGCGCGCGCGCGCGCGGAGACAA 120  
Db 264 GTCCGGCAGCGGTAGCAGTGGCGGAGCAGCAGCGCGCGCGCGCGCGGAGACAA 323  
QY 121 AGCATGGAGCAGCAGCAAGATGGAGAAAGCGGACTCTCGGGAATCTTCATTTCACT 180  
Db 324 AGCATGGAGCAGCAGCAAGATGGAGAAAGCGGACTCTCAGGAATCTTCATTTCACT 383  
QY 181 GGTTCATGGTATGATTCATTTGCGGCTCTGACATCTCTAGTCTGCTTGGTTGATC 240  
Db 384 GGTTCATGGTATGATTCATTTGCGGCTCTGACATCTCTAGTCTGCTTGGTTGATC 443  
QY 241 TTGTTGATGAGGAGTCTAGGAAACTAGGATCTATGATCTGATGGTGGAG 300  
Db 444 TTGTTGATGAGGAGTCTAGGAAACTAGGATCTATGATCTGATGGTGGAG 503  
QY 301 ATTTTCATGGTATGATGCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTCAG 357  
Db 504 ATTTTCATGGTATGATGCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTCAG 560

RESULT 69

CN368828 548 bp mRNA linear EST 16-MAY-2004  
LOCUS 17000424344539 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
DEFINITION CN368828  
ACCESSION CN368828  
VERSION CN368828.1 GI:47368762  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 548)  
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
Lebkowski, J. and Stanton, L.W.  
TITLE Transcription characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)  
COMMENT Contact: Brandenberger R.  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 548 Std Error: 0.00.  
Location/Qualifiers  
1. 548  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/note="Oligo dt primed, full-length enriched cDNA library

FEATURES source

Matches 408; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 153 GGACTCTCGGAACTTCATTTTCACGTGTTTATGTTGATTCATTCATTCGTCGTCG 212  
Db 125 GGACTCTCGGAACTTCATTTTCACGTGTTTATGTTGATTCATTCATTCGTCGTCG 184  
QY 213 ACATCTGTAGTCTGCTGTTGGTTGATCTTGTGACTATGAGAACTTCTAGAAAAC 272  
Db 185 ACATCTGTAGTCTGCTGTTGGTTGATCTTGTGACTATGAGAACTTCTAGAAAAC 244  
QY 273 GGATCTATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 332  
Db 245 GGATCTATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 304  
QY 333 GGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCAC 392  
Db 305 GGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCAC 364  
QY 393 ACTGAGCCGAGGAGGAGTCTCTGAGGAGCAAGCCAGCAATATCGAAGATGAAGCA 452  
Db 365 ACTGAGCCGAGGAGGAGTCTCTGAGGAGCAAGCCAGCAATATCGAAGATGAAGCA 424  
QY 453 AAAGAACTTCACTGCTCCCTTCTCCATGAATGTTACACGAGCAACATGTTGAGGAGAA 512  
Db 425 AAAGAACTTCACTGCTCCCTTCTCCATGAATGTTACACGAGCAACATGTTGAGGAGAA 484  
QY 513 GACTTGCACAAAGAGATGAGCCACAGGAGAAACACACAGAGGATGA 562  
Db 485 GACTTGCACAAAGAGATGAGCCACAGGAGAAACACACAGAGGATGA 534

RESULT 68

CN482623 560 bp mRNA linear EST 26-APR-2004  
LOCUS hw22f04.y1 Human primary human ocular pericytes. Unamplified (hw)  
DEFINITION CN482623  
ACCESSION CN482623  
VERSION CN482623.1 GI:46564127  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 560)  
AUTHORS Tsai, J.Y. and Wistow, G.  
TITLE Expressed sequence tag analysis of cultured primary human ocular  
pericytes  
JOURNAL Unpublished (2004)  
COMMENT Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 22 row: f column: 04  
Seq primer: M13RPI reverse primer (ABI).  
Location/Qualifiers  
1. 560  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="hw22f04"  
/cell\_type="pericytes"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human primary human ocular pericytes.  
Unamplified (hw)"  
/note="Organ: Eye; Vector: pSPORT1; RNA was extracted from  
primary human pericytes in culture. A directionally cloned  
cDNA library in the pSPORT1 vector (Invitrogen) was  
constructed at Bioserve Biotechnology (Laurel MD)  
essentially following the protocols of the SuperScript

FEATURES source





Site 2: XhoI: The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 12.7%; Score 294; DB 4; Length 536;  
Best Local Similarity 100.0%; Pred. No. 8e-146;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 508 GAGAAGACTTCCACACAGAAGATGACCCACAGGAGAACACACACACAGAGGATGATGAGT 567  
Db 1 GAGAAGACTTCCACACAGAAGATGACCCACAGGAGAACACACACAGAGGATGATGAGT 60

Qy 568 TTCTTATGGGACTGATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 627  
Db 61 TTCTTATGGGACTGATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 120

Qy 628 AAGAACCAGCAGATGATTTACCACTGGGAAGAGACAGTTTCAACAGCTGTAATCAGGATA 687  
Db 121 AAGAACCAGCAGATGATTTACCACTGGGAAGAGACAGTTTCAACAGCTGTAATCAGGATA 180

Qy 688 TGAAGACAGATGCTGACGAGGAGGAATCCAGATCCAGTGAACCCAGTGTAGAGATG 747  
Db 181 TGAAGACAGATGCTGACGAGGAGGAATCCAGATCCAGTGAACCCAGTGTAGAGATG 240

Qy 748 AAGAGTTGCACCATGATACAGATGATGTAACATACCAAGTCTATGAGGAACAAG 801  
Db 241 AAGAGTTGCACCATGATACAGATGATGTAACATACCAAGTCTATGAGGAACAAG 294

RESULT 72

AV751514 622 bp mRNA linear EST 19-OCT-2000  
LOCUS AV751514 NPD Homo sapiens cDNA clone NPDXB01 5', mRNA sequence.  
DEFINITION AV751514  
ACCESSION AV751514  
VERSION AV751514.1 GI:10909362  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 622)  
Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q., Han,Z., Chen,Z., Hu,R. and Chen,J.  
Homo sapiens NPD library cDNA clones  
Unpublished (2000)  
Contact: Qinghua Zhang  
Shanghai Institute of Endocrinology, Rui-Jin Hospital  
197 Rui-Jin II Road, Shanghai 200025, P. R. China  
Tel: 86-21-64370045 (ex.663332)  
Fax: 86-21-64743206  
Email: mbsni@ms.sh.cn  
This clone is available at Shanghai Hematology Institute in Shanghai.  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.  
Location/Qualifiers  
1. .622  
/organism="Homo sapiens"  
/mol\_type="mRNA"

FEATURES  
source

/db\_xref="taxon:9606"  
/clone="NPDXB01"  
/tissue\_type="pituitary"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_lib="NPD"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN

Query Match 12.7%; Score 294; DB 1; Length 622;  
Best Local Similarity 100.0%; Pred. No. 8.2e-146;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 653 GGAAGAGACAGATTTCCACAGACTGTAAATCAGGATATGGAAGAGATGATCTCTGAGCAGGA 712  
Db 90 GGAAGAGACAGATTTCCACAGACTGTAAATCAGGATATGGAAGAGATGATCTCTGAGCAGGA 149

Qy 713 AAATCCAGATTCAGTGAACCCAGTAGTAGTAAGATGAAGATTCACCATGATACAGATGA 772  
Db 150 AAATCCAGATTCAGTGAACCCAGTAGTAGTAAGATGAAGATTCACCATGATACAGATGA 209

Qy 773 TGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGGAT 832  
Db 210 TGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGGAT 269

Qy 833 AGAAATCAGACAGAGTAGTAACCTCTCCCTCAGGATATATCTGTAGAAAGATTCACAGTAAT 892  
Db 270 AGAAATCAGACAGAGTAGTAACCTCTCCCTCAGGATATATCTGTAGAAAGATTCACAGTAAT 329

Qy 893 TGTAGAAAGATGAAGATTTTCTCTGGAAGAACAGCAGGAGGAGTACCACCCAGA 946  
Db 330 TGTAGAAAGATGAAGATTTTCTCTGGAAGAACAGCAGGAGGAGTACCACCCAGA 383

RESULT 73

CD512083 862 bp mRNA linear EST 06-JUN-2003  
LOCUS AGENCOURT\_14354806 NIH\_MGC\_186 Homo sapiens cDNA clone  
DEFINITION IMAGE:30406369 5', mRNA sequence.  
ACCESSION CD512083  
VERSION CD512083.1 GI:31443801  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 862)  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cagpos-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDCMI95 row: a column: 02  
High quality sequence stop: 564.

FEATURES  
1. .862  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30406369"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_186"

/note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI (ggccatagcc); Site 2: SfiI (ggcgctggcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a pooled samples of tissues from Skin, meninges, duramater, pia mater and choroid plexus. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGAGCGCCGACATG-dT(30)BN-3' (where B = A, C, G and N = A, G, C or T). Average insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library"

## ORIGIN

Query Match 12.7%; Score 294; DB 6; Length 862;  
Best Local Similarity 100.0%; Pred No. 8.7e-146; Indels 0; Gaps 0;  
Matches 294; Conservative 0; Mismatches 0;

QY 653 GGAAGAGACAGTTTCAAGACTGTAAATCAGGATATGGAAGATGATGTCTGAGCAGA 712  
DB 3 GGAAGAGACAGTTTCAAGACTGTAAATCAGGATATGGAAGATGATGTCTGAGCAGA 62  
QY 713 AAATCCAGATTCCAGTGAACACAGTAGTAGAAGATGAAAGATTGCACCATGATACAGATGA 772  
DB 63 AAATCCAGATTCCAGTGAACACAGTAGTAGAAGATGAAAGATTGCACCATGATACAGATGA 122  
QY 773 TGTACATACCAAGTCTATGAGCAACAGCAGTATATCAACCTCTAGAAATCAAGGGAT 832  
DB 123 TGTACATACCAAGTCTATGAGCAACAGCAGTATATCAACCTCTAGAAATCAAGGGAT 182  
QY 833 AGAATACACAGTAACTCTCTCCCTCAGGATATCTCTAGAGATTACAGGTAAT 892  
DB 183 AGAATACACAGTAACTCTCTCCCTCAGGATATCTCTAGAGATTACAGGTAAT 242  
QY 893 TGTAGAAGATTAAGCAATTTTCTCTGTGAAGAACAGCAGGAAGTACCACCA 946  
DB 243 TGTAGAAGATTAAGCAATTTTCTCTGTGAAGAACAGCAGGAAGTACCACCA 296

## RESULT 74

BM462707  
LOCUS BM462707 1010 bp mRNA linear EST 05-FEB-2002  
DEFINITION AGENCOURT 6427239 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5540329  
5', mRNA sequence.

ACCESSION BM462707  
VERSION BM462707.1 GI:18511747

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1010)  
NIH-MGC <http://mgi.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
UNPUBLISHED (1999)

TITLE Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapbe@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTT

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LAM12236 row: e column: 02

High quality sequence stop: 579.

Location/Qualifiers

1. 1010

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

## FEATURES

source

/clone="IMAGE:5540329"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: Skin; Vector: pCMV-Sport6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

## ORIGIN

Query Match 12.4%; Score 288; DB 4; Length 1010;  
Best Local Similarity 99.5%; Pred No. 1.5e-142;  
Matches 388; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATTTGCCCGCAGCGTAAAGATGCCAAGAGCAGCGCCGACAGCAGCAGCG 60  
DB 206 CGGACCGTGCATTTGCCCGCAGCGTAAAGATGCCAAGAGCAGCGCCGACAGCAGCAGCG 265  
QY 61 GCTCCGCGAGCGGTAGCAGAGTGGCGGCGAGCAGCAGCCCGCGGCGGAGAGACAA 120  
DB 266 GCTCCGCGAGCGGTAGCAGAGTGGCGGCGAGCAGCAGCCCGCGGCGGAGAGACAA 325  
QY 121 AGCATGGAGGACACAAGAAATGGGAGGAAAGCGGAGCTCTCGGGAATCTTCACTTCCACGT 180  
DB 326 AGCATGGAGGACACAAGAAATGGGAGGAAAGCGGAGCTCTCGGGAATCTTCACTTCCACGT 385  
QY 181 GGTATTATGTTGATTGCTGGCGCTCTGGACATCTGTAGCTGCTGTTGGTTGATC 240  
DB 386 GGTATTATGTTGATTGCTGGCGCTCTGGACATCTGTAGCTGCTGTTGGTTGATC 445  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTATAGAACTATATGCTGATGCTGATGGAG 300  
DB 446 TTGTTGACTATGAGGAAGTTCTAGGAAACTATAGAACTATATGCTGATGCTGATGGAG 505  
QY 301 ATTTTGTATGTTGATGATGTCACAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC 360  
DB 506 ATTTTGTATGTTGATGATGTCACAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC 565  
QY 361 CAGCAGTCCCGCCAGAGAGAGGCTGAGCCAC 390  
DB 566 CAGCAGTCCCGCCAGAGAGAGGCTGAGCCAC 595

## RESULT 75

CN368820

LOCUS CN368820

DEFINITION 17000533684042 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CN368820

VERSION CN368820.1 GI:47368754

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 767)

AUTHORS Brandenberger R., Wei, H., Zhang S., Lei S., Murage J., Fisk G.J.,

Li Y., Xu, C., Pang, R., Guebler, K., Rao, M.S., Mandalam, R.,

Lebkowski, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)

Contact: Brandenberger R

Regenerative Medicine

Genon Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbrandenberger@genon.com

Insert Length: 767 Std Error: 0.00.

Location/Qualifiers

1. 767

/organism="Homo sapiens"

/mol\_type="mRNA"

/db xref="taxon:9606"  
/tissue type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/notes="oligo dt primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

ORIGIN

Query Match 12.2%; Score 283; DB 7; Length 767;  
Best Local Similarity 99.7%; Pred. No. 7.1e-140;  
Matches 333; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATGCCCGCAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 180 CGGACCGTGCATGCCCGCAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 239  
Qy 61 GCTCCGGCAGCGGTAGCACAGAGTCCGGGCGAGCAGCAGCCCGGGGCCCGGAGAGAGACAA 120  
Db 240 GCTCCGGCAGCGGTAGCACAGAGTCCGGGCGAGCAGCAGCCCGGGGCCCGGAGAGAGACAA 299  
Qy 121 AGCATGGAGGACACAAGAAATGGAGGAAAGCCGACTCTCGGGAACCTTCATTCTTACGT 180  
Db 300 AGCATGGAGGACACAAGAAATGGAGGAAAGCCGACTCTCAGGAACTTCATTCTTACGT 359  
Qy 181 GGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATC 240  
Db 360 GGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATC 419  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGCTGATG 300  
Db 420 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGCTGATG 479  
Qy 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 334  
Db 480 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 513

Search completed: March 25, 2005, 22:28:27  
Job time : 6796 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 14:29:03 ; Search time 1157 Seconds  
(without alignments)  
11890.647 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1980s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2000s:\*  
5: Geneseqn2000s:\*  
6: Geneseqn2000s:\*  
7: Geneseqn2000s:\*  
8: Geneseqn2000s:\*  
9: Geneseqn2000s:\*  
10: Geneseqn2000s:\*  
11: Geneseqn2000s:\*  
12: Geneseqn2000s:\*  
13: Geneseqn2000s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2324	100.0	2324	4	Aaf89811 cDNA enco
2	2324	100.0	2324	8	Abs57868 Human cDN
3	2324	100.0	2324	9	Ada00640 cDNA enco
4	2322	95.6	2452	13	Adp23914 PRO polyp
5	2322	95.6	2452	13	Adp23914 PRO polyp
6	2171	93.4	5358	10	Adj56257 Bovine CD
7	895	38.5	2680	6	Abx04178 Human mRN
8	895	38.5	2680	13	Adp24208
9	840	27.5	2442	2	Aaz23609 PRO polyp
10	840	27.5	2442	6	Abk39743 cDNA enco
11	840	27.5	2442	8	Acad2072 Human lun
12	840	27.5	2442	8	Acad2072 Human lun
13	640	27.5	2442	10	Adh47310 Human lun
14	640	27.5	2442	13	Adj21229 Human lun
15	615	26.5	3110	11	Acn88788 Breast ca
16	502	21.6	2648	10	AcA92448
17	459	19.8	502	9	ACH47067
18	359	15.4	660	6	ABQ59507
19	249	10.7	495	11	ADT94758
20	104	4.5	161	12	ACH93769 Human gen

C	21	104	4.5	592	12	ACH80069
	22	72	3.1	1382	2	AAX22124
	23	72	3.1	1382	8	ADA39640 Human sec
	24	72	3.1	1382	8	ACC50349 Human sec
	25	72	3.1	1382	10	ADA55835
	26	60	2.6	60	6	ABN38322
	27	59	2.5	2208	11	ADN95564
	28	59	2.5	2208	13	ADN95564 Human BSC
	29	59	2.5	2208	13	ADN95564 Human NF-
	30	51	2.2	583	6	ABQ59571
	31	27	1.2	27	13	ADR97461
	32	27	1.0	23	13	ADR97432
	33	22	0.9	855	4	ABL25903
	34	22	0.9	1933	4	ABL19239
	35	22	0.9	2227	4	ABL09457
	36	22	0.9	2855	4	ABL25902
	37	22	0.9	3705	4	ABL23387
	38	22	0.9	3933	4	ABL19238
	39	22	0.9	3974	4	ABL09456
	40	22	0.9	5903	4	ABL23386
	41	21	0.9	21	13	ADR97567
	42	21	0.9	1845	3	AZ98400
	43	21	0.9	1845	6	ABK40732
	44	21	0.9	4011	2	AAX60124
	45	20	0.9	20	4	AAR9814
	46	20	0.9	20	4	AAR9813
	47	20	0.9	20	4	AAR9812
	48	20	0.9	20	8	ABS57871
	49	20	0.9	20	8	ABS57870
	50	20	0.9	20	8	ABS57869
	51	20	0.9	21	13	ADR97507
	52	20	0.9	550	12	ACH75949
	53	20	0.9	786	11	ACH96138
	54	20	0.9	849	2	AAT62523
	55	20	0.9	900	5	ABA19178
	56	20	0.9	908	3	AAC58389
	57	20	0.9	936	6	AAL49917
	58	20	0.9	984	4	ABL27131
	59	20	0.9	1038	13	ACN42014
	60	20	0.9	1039	13	ACN42013
	61	20	0.9	1070	4	ABL16775
	62	20	0.9	1088	13	ACN42012
	63	20	0.9	1089	13	ACN42011
	64	20	0.9	1109	2	AAX24880
	65	20	0.9	1230	3	AAAI3712
	66	20	0.9	1245	10	ADM28660
	67	20	0.9	1245	12	ADM93405
	68	20	0.9	1418	10	ACC79031
	69	20	0.9	1565	4	ABL04193
	70	20	0.9	1891	12	ADM47862
	71	20	0.9	2000	8	ADA71941
	72	20	0.9	2000	12	ADJ41021
	73	20	0.9	2370	4	AAH18422
	74	20	0.9	2370	5	AAS72135
	75	20	0.9	2803	12	ADG31190
	76	20	0.9	3008	4	ABL15949
	77	20	0.9	3144	4	ABL27130
	78	20	0.9	3224	4	ABL16774
	79	20	0.9	3369	13	ADR07848
	80	20	0.9	3689	5	AAS83545
	81	20	0.9	4021	4	ABL04192
	82	20	0.9	4275	4	ABL05197
	83	20	0.9	5709	4	ABL15948
	84	20	0.9	6959	4	ABL05196
	85	20	0.9	8772	4	AAS11678
	86	20	0.9	8772	12	ADO29841
	87	20	0.9	8871	4	AAS11677
	88	20	0.9	9321	6	ABK15177
	89	20	0.9	9613	4	ABL02636
	90	20	0.9	10531	8	ABZ42868
	91	20	0.9	10694	11	ADL22585
	92	20	0.9	17951	4	ABL11844
	93	20	0.9	24370	10	ADC86478

94	20	0.9	28564	10	ADD45153	Human gen
95	20	0.9	48037	4	AAK84729	Human imm
96	20	0.9	48037	4	AAK85983	Human imm
97	20	0.9	48045	4	AAK84730	Human imm
98	20	0.9	48045	4	AAK85984	Human imm
99	20	0.9	52677	11	ACNA43934	Human gen
c 100	20	0.9	159400	6	ABQ88126	Human ost

## ALIGNMENTS

RESULT 1	
AAF89811	AAF89811 standard; cDNA; 2324 BP.
ID	AAF89811 standard; cDNA; 2324 BP.
XX	
XX	AAF89811;
AC	
XX	
XX	23-JUL-2001 (first entry)
DT	
DT	
XX	
XX	cDNA encoding human aspartyl (asparaginyl) beta-hydroxylase (HAAH).
DE	
XX	
XX	Epidermal growth factor-like domain; EGF-like domain; cancer;
KW	human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour; ss.
KW	
XX	
XX	Homo sapiens.
OS	

AA  
FN WO200135102-A2.  
XX  
XX  
XX 17-MAY-2001.  
XX  
XX 08-NOV-2000; 2000WO-US030738.  
XX  
XX 08-NOV-1999; 99US-00436184.  
XX  
XX (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.  
PA  
XX  
XX Wanda JR, De La Monte SM, Ince N, Carlson RI;  
PI  
XX WPI; 2001-329171/34.  
XX P-PSDB: AAB83919.  
DR

The present sequence encodes a human aspartyl (asparaginyl) beta-hydroxylase (HAAH) enzyme. Epidermal growth factor (EGF)-like domains of polypeptides are hydroxylated by HAAH enzymes. HAAH is used in the method of the invention. The specification describes a method for diagnosing a malignant neoplasm in a mammal. The method comprises contacting a body fluid with an antibody which binds to HAAH polypeptide under complex forming conditions, and detecting the antigen-antibody complex. The method is useful for diagnosing and prognosing a malignant neoplasm in a bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood, serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal, where the neoplasm is derived from endodermal tissue and is selected from colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic agent, are useful for killing tumour cells.

SQ Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 2324; DB 4; Length 2324;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGGACCGTGAATGGCCCAAGCTAAGATCCCAAGAGCAGCGGCAACAGCACGACGCG	60
Db	1	CGGACCGTGAATGGCCCAAGCTAAGATCCCAAGAGCAGCGGCAACAGCACGACGCG	60
Qy	61	GCTCCGGCAGCGGTAGCAGAGTCGGGCGAGCAGACAGCCCCCGGGGCCGAGAGACAA	120
Db	61	GCTCCGGCAGCGGTAGCAGAGTCGGGCGAGCAGACAGCCCCCGGGGCCGAGAGACAA	120
Qy	121	AGCATGGAGACACAAGAAATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCATCTTCACGT	180
Db	121	AGCATGGAGACACAAGAAATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCATCTTCACGT	180
Qy	181	GGTTTATGTGATTCGAAATGCTGGGCGCTGGACATCTCTAGCTGTCTGTTGTTGATC	240
Db	181	GGTTTATGTGATTCGAAATGCTGGGCGCTGGACATCTCTAGCTGTCTGTTGTTGATC	240
Qy	241	TTGTTCACTATGAGGAAGTTTCTAGGAAAACTAGGAATCTATGATGCTGATGTTGATGGAG	300
Db	241	TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGTTGATGGAG	300
Qy	301	ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCCAGAGC	360
Db	301	ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCCAGAGC	360
Qy	361	CACAGTCCCGCAGAAAGAGGCTGAGCCACACACTCAGCCCCGAGGACGAGTTCCTGTGG	420
Db	361	CACAGTCCCGCAGAAAGAGGCTGAGCCACACACTCAGCCCCGAGGACGAGTTCCTGTGG	420
Qy	421	AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAATTCAGTCCCTCTCCATG	480
Db	421	AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAATTCAGTCCCTCTCCATG	480
Qy	481	AAATGTTACACGACAGAACATGTTGAGGAGAGACATTCGAAACAAGAAAGATGGAACCCACAG	540
Db	481	AAATGTTACACGACAGAACATGTTGAGGAGAGACATTCGAAACAAGAAAGATGGAACCCACAG	540
Qy	541	GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTTTG	600
Db	541	GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTTTG	600
Qy	601	AGACCTTGGAAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCAAGTGGAGAGA	660
Db	601	AGACCTTGGAAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCAAGTGGAGAGA	660
Qy	661	CAGTTTCACAAGACTGTAAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAAAATCCAG	720
Db	661	CAGTTTCACAAGACTGTAAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAAAATCCAG	720
Qy	721	ATTCAGTGAAACCAAGTAGTAGAAGATGAAGATTCACATGATACAGATGATGTAACAT	780
Db	721	ATTCAGTGAAACCAAGTAGTAGAAGATGAAGATTCACATGATACAGATGATGTAACAT	780
Qy	781	ACCAAGTCTATGAGGAACAGCAGTATATGAACTCTAGAAAAATGAAGGATAGAAATCA	840
Db	781	ACCAAGTCTATGAGGAACAGCAGTATATGAACTCTAGAAAAATGAAGGATAGAAATCA	840
Qy	841	CAGAAGTAACTGTCTCCCTCAGGATAATCCTGTAGAAGATTCACAGGTAATTTGTAAG	900
Db	841	CAGAAGTAACTGTCTCCCTCAGGATAATCCTGTAGAAGATTCACAGGTAATTTGTAAG	900
Qy	901	AAGTAAGCATTTTCTGTGTGAAGACAGCAGGAAGTACCACAGAAACAAATAGAAAAA	960
Db	901	AAGTAAGCATTTTCTGTGTGAAGACAGCAGGAAGTACCACAGAAACAAATAGAAAAA	960
Qy	961	CAGATGATCCAGAACAAAAAGTAAAGAAAAAGAACCCCTAACTTTTAAATAAAT	1020
Db	961	CAGATGATCCAGAACAAAAAGTAAAGAAAAAGAACCCCTAACTTTTAAATAAAT	1020
Qy	1021	TTGATAAGACTATTAAGCTGAACCTTCATGCTGCAGAAAAAATCCGTAAAGGGGAAAAA	1080
Db	1021	TTGATAAGACTATTAAGCTGAACCTTCATGCTGCAGAAAAAATCCGTAAAGGGGAAAAA	1080

1081 TTGAGGAGCAGTGAATGCTATTAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
1081 TTGAGGAGCAGTGAATGCTATTAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGGAGGAGTATGAGG 1200  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGGAGGAGTATGAGG 1200  
1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCAGCCTACCTGATGTCCTGGCAG 1260  
1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCAGCCTACCTGATGTCCTGGCAG 1260  
1261 ACTGCTGAAGCTGAGTTTGAAGCGTCTGAGAGGCAACAAATTTCTAGGTCATATGA 1320  
1261 ACTGCTGAAGCTGAGTTTGAAGCGTCTGAGAGGCAACAAATTTCTAGGTCATATGA 1320  
1321 GAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCACTATTTCCCAATGATACTTCTTAA 1380  
1321 GAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCACTATTTCCCAATGATACTTCTTAA 1380  
1381 AAAATGACCTTGGCGTGGATACCTCTGTATAGGAGATATGACAAATGCCAAGAGTTT 1440  
1381 AAAATGACCTTGGCGTGGATACCTCTGTATAGGAGATATGACAAATGCCAAGAGTTT 1440  
1441 ATGAAGAGGTCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
1441 ATGAAGAGGTCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
1501 TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAAGTAAGAT 1560  
1501 TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAAGTAAGAT 1560  
1561 CCGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGATGCCATCGAGA 1620  
1561 CCGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGATGCCATCGAGA 1620  
1621 GGGTTGGGACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAAGAGAGGACATTTG 1680  
1621 GGGTTGGGACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAAGAGAGGACATTTG 1680  
1681 CATCTGTCTGCAACCGTCACTACAATGTGAATGAGTGAATGAGTGAATGAGTGAATG 1740  
1681 CATCTGTCTGCAACCGTCACTACAATGTGAATGAGTGAATGAGTGAATGAGTGAATG 1740  
1741 CCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800  
1741 CCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800  
1801 TCCGAGATGAAGGCTTGGCAGTATGATGATTAAGCCAAAGGCTCTCTTCCCTGAGGATG 1860  
1801 TCCGAGATGAAGGCTTGGCAGTATGATGATTAAGCCAAAGGCTCTCTTCCCTGAGGATG 1860  
1861 AAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920  
1861 AAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920  
1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAAAGTTCCCGGAGA 1980  
1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAAAGTTCCCGGAGA 1980  
1981 CAACAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040  
1981 CAACAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040  
2041 GGGCGCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
2041 GGGCGCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
2101 AGGAAGGCTGCAAGATTCGATGTGCAACGAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 2160  
2101 AGGAAGGCTGCAAGATTCGATGTGCAACGAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 2160  
2161 TCATCTTTTGTACTCTCTTTGAGCAGGAGTATGGCAGGATGCTCATCTTTCCGGCTGA 2220

Db 2161 TCATCTTTTGTACTCTCTTTGAGCAGGATATGGCAGATGCTCATCTTTCCGGCTGA 2220  
Qy 2221 TATTCATCGTGGATGTGTGCGATCCGGAATCTGACACCAAGAGAGGAGGAGGAGGAGG 2280  
Db 2221 TATTCATCGTGGATGTGTGCGATCCGGAATCTGACACCAAGAGAGGAGGAGGAGGAGG 2280  
Qy 2281 CAATTTAGCATGAATTCATCAAGCTTGGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 2  
ABS57868  
ID ABS57868 standard; CDNA; 2324 BP.  
XX  
AC ABS57868;  
XX  
DT 06-FEB-2003 (first entry)  
XX  
DE Human cDNA encoding aspartyl (asparaginyl) beta-hydroxylase, HAAH.  
XX  
KW Human; ss; gene; aspartyl (asparaginyl) beta-hydroxylase; HAAH;  
KW cytototoxic; immunostimulant; antibody; neoplasm; tumour; PB50; 5C7;  
KW 19B; brain tumour; glioma; glioblastoma; astrocytoma; haemangioma;  
KW pancreatic carcinoma; colon cancer; breast cancer; pancreatic cancer;  
KW liver cancer; cancer of the bile ducts; primary malignant CNS neoplasm;  
KW metastatic CNS neoplasm.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT 12..2288  
FT /\*tag= a  
FT /product= "HAAH"  
XX  
US2002110559-A1.  
XX  
XX 15-AUG-2002.  
XX  
XX 17-MAY-2001; 2001US-00859604.  
XX  
XX 08-NOV-1999; 99US-00436184.  
XX  
XX (WAND/) WANDS J R.  
XX (DMON/) DE LA MONTE S M.  
XX (DEUT/) DEUTCH A H.  
XX (GHAN/) GHANBARI H A.  
XX  
XX Wands JR, De La Monte SM, Deutch AH, Ghanbari HA;  
XX  
XX WPI; 2003-066676/06.  
XX P-PSDB; ABG72365.  
XX  
XX Diagnosing neoplasm and inhibiting tumor growth, by contacting tissue of  
XX mammal with detectably-labeled antibody which binds to human aspartyl  
XX (asparaginyl) beta-hydroxylase.  
XX  
XX Example 1; Page 4-5; 34pp; English.  
XX  
XX The invention relates to diagnosing a neoplasm and inhibiting tumour  
XX growth in a mammal, using an antibody that binds to human aspartyl  
XX (asparaginyl) beta-hydroxylase (HAAH). Diagnosing a neoplasm comprises  
XX contacting a tissue with a detectably-labelled antibody where an increase  
XX in level of antibody binding at tissue site compared to the level of  
XX binding to normal non-neoplastic tissue indicates the presence of a  
XX neoplasm at the tissue site. Inhibiting tumour growth in mammal involves  
XX administering the antibody conjugated to a cytotoxic agent to a mammal.  
XX Also included are a method of conferring an immune response to a tumour  
XX cell in a mammal, by administering the antibody, a method of inducing an  
XX HAAH-specific immune response in a mammal by administering to the mammal  
XX an HAAH polypeptide (or a polynucleotide composition encoding the  
XX polypeptide, or its degenerate variant), a fragment of HAAH comprising an

CC extracellular domain and lacking a cytoplasmic domain of HAAH, an  
CC antibody or its fragment which binds to HAAH (where the antibody is FB50,  
CC 86A, 5C7 or 19B), a hybridoma cell line chosen from hybridoma FB50,  
CC HA386A, HA15CA and HA219B, and a fragment of HAAH which lacks enzymatic  
CC activity or alpha-ketoglutarate binding domain and epidermal growth  
CC factor (EGF)-like domain. The methods are useful for diagnosing neoplasm  
CC in a mammal, inhibiting tumour growth in a mammal, conferring an immune  
CC response to a brain tumour cell (e.g. glioma, glioblastoma, astrocytoma  
CC or haemangioma) in a mammal, for conferring immune response to a  
CC pancreatic carcinoma cell and for inducing a HAAH-specific immune  
CC response in a mammal. The method is useful for diagnosing malignant  
CC neoplasms derived from endodermal tissue, e.g. colon cancer, breast  
CC cancer, pancreatic cancer, liver cancer and cancer of the bile ducts. The  
CC method is also useful for diagnosing neoplasms of central nervous system  
CC (CNS) e.g. primary malignant CNS neoplasms, and for diagnosing brain  
CC cell origin and metastatic CNS neoplasms, and for diagnosing brain  
CC tumours e.g. glioma, glioblastoma, astrocytoma or haemangioma. The  
CC present sequence encodes human HAAH  
XX  
SQ Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2324; DB 8; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATCGCCAGCGTAAGATGCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 1 CGGACCGTGCATCGCCAGCGTAAGATGCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
QY 61 GCTCCGGCAGCGGTAGCAGTAGTGGCGCAGCAGCAGCGCCCGCGGAGAGACAA 120  
DB 61 GCTCCGGCAGCGGTAGCAGTAGTGGCGCAGCAGCAGCGCCCGCGGAGAGACAA 120  
QY 121 AGCATGGAGGACACAGAAATGGGAGGAAAGGGGACTCTCGGGAACCTTCATCTTCAGT 180  
DB 121 AGCATGGAGGACACAGAAATGGGAGGAAAGGGGACTCTCGGGAACCTTCATCTTCAGT 180  
QY 181 GGTATTATCGTATTCATTCGTGGCGCTGTGACATCTGTAGCTGTCTGTTGGTTGATC 240  
DB 181 GGTATTATCGTATTCATTCGTGGCGCTGTGACATCTGTAGCTGTCTGTTGGTTGATC 240  
QY 241 TTGTTGACTATGAGAAAGTTCTAGAAATCTAGGAATCTATGATCTGATGGTGTGAG 300  
DB 241 TTGTTGACTATGAGAAAGTTCTAGAAATCTAGGAATCTATGATCTGATGGTGTGAG 300  
QY 301 ATTTTGTATGTGATGATGCCAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
DB 301 ATTTTGTATGTGATGATGCCAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
QY 361 CAGCAGTCCCGCAGAAAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGTG 420  
DB 361 CAGCAGTCCCGCAGAAAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGTG 420  
QY 421 AGGCAGAACCCAGAAATTCGAAGATGAAGCAAAAGACAAATTCAGTCCCTTCTCCATG 480  
DB 421 AGGCAGAACCCAGAAATTCGAAGATGAAGCAAAAGACAAATTCAGTCCCTTCTCCATG 480  
QY 481 AAATGGTACCCAGCAACATGTTGAGGAGAGAGACTTGCAACAAGAGATGAGCCACAG 540  
DB 481 AAATGGTACCCAGCAACATGTTGAGGAGAGAGACTTGCAACAAGAGATGAGCCACAG 540  
QY 541 GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTTG 600  
DB 541 GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTTG 600  
QY 601 AGACCTGGAACTGAGATCTCATGAGAAACCGGACATAGTTACCACTGGAGAGA 660  
DB 601 AGACCTGGAACTGAGATCTCATGAGAAACCGGACATAGTTACCACTGGAGAGA 660  
QY 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 720  
DB 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 720

QY 721 ATTCCAGTGAAACAGTAGTAGAAGATGAAGATTGCAACCATGATACAGATGATTAACAT 780  
DB 721 ATTCCAGTGAAACAGTAGTAGAAGATGAAGATTGCAACCATGATACAGATGATTAACAT 780  
QY 781 ACCNAGTCTTACGAGCAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
DB 781 ACCNAGTCTTACGAGCAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
QY 841 CAGAACTAACTGCTCCCTCGAGGATAATCTCTGTAGAAGATTACACAGGTAATTTGTAGAAG 900  
DB 841 CAGAACTAACTGCTCCCTCGAGGATAATCTCTGTAGAAGATTACACAGGTAATTTGTAGAAG 900  
QY 901 AAGTAAGCAATTTTCTGTGGAAGAACACAGCAGGAAGTACACAGAAACAAATAGAAAAA 960  
DB 901 AAGTAAGCAATTTTCTGTGGAAGAACACAGCAGGAAGTACACAGAAACAAATAGAAAAA 960  
QY 961 CAGATGATCCAGAACAAAGCAAAAGTTAAAGAAAAAGAGCTAACTTTTAAATAAAT 1020  
DB 961 CAGATGATCCAGAACAAAGCAAAAGTTAAAGAAAAAGAGCTAACTTTTAAATAAAT 1020  
QY 1021 TTGATAAGACTATTTAAAGCTGAACTTTGATGCTGCAGAAAAAATCTCCGTAAAAAGGGAAAAA 1080  
DB 1021 TTGATAAGACTATTTAAAGCTGAACTTTGATGCTGCAGAAAAAATCTCCGTAAAAAGGGAAAAA 1080  
QY 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGAAATACCTCTCAGAGTCCACGAG 1140  
DB 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGAAATACCTCTCAGAGTCCACGAG 1140  
QY 1141 CAAGATATGGAAGGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGAGAGTAATGAGG 1200  
DB 1141 CAAGATATGGAAGGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGAGAGTAATGAGG 1200  
QY 1201 TGCTACGTGAGCCATCGAGACCTTACAGAGGTGGCGAGCCTACCTGATGTCCTGTCAG 1260  
DB 1201 TGCTACGTGAGCCATCGAGACCTTACAGAGGTGGCGAGCCTACCTGATGTCCTGTCAG 1260  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320  
DB 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320  
QY 1321 GAGGTTCCCTGCTTACCTGCGAGAGTATGTTCAACTATTTTCCCAATGATACCTTCTTAA 1380  
DB 1321 GAGGTTCCCTGCTTACCTGCGAGAGTATGTTCAACTATTTTCCCAATGATACCTTCTTAA 1380  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATGCAATTCGAAAGAAAGTTT 1440  
DB 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATGCAATTCGAAAGAAAGTTT 1440  
QY 1441 ATGAAGAGGTGCTGAGGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
DB 1441 ATGAAGAGGTGCTGAGGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
QY 1501 TCCTGAAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGAGAGAAATAGAA 1560  
DB 1501 TCCTGAAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGAGAGAAATAGAA 1560  
QY 1561 CCGGAGATCTTGGCAGTGAATGGGATTTTATTTCCACCTGGGGATGCCATGCGAGA 1620  
DB 1561 CCGGAGATCTTGGCAGTGAATGGGATTTTATTTCCACCTGGGGATGCCATGCGAGA 1620  
QY 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGTTGGGCACAAAGAGAGACACTTTG 1680  
DB 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGTTGGGCACAAAGAGAGACACTTTG 1680  
QY 1681 CATCTGCTGCAACCGCTCACTCTACATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1740  
DB 1681 CATCTGCTGCAACCGCTCACTCTACATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1740  
QY 1741 CCCCAAAAGAAACGGGCTTACAGAGTTAGTAAAGTCTTTTAAAGAACTGGAAGTTAA 1800  
DB 1741 CCCCAAAAGAAACGGGCTTACAGAGTTAGTAAAGTCTTTTAAAGAACTGGAAGTTAA 1800  
QY 1801 TCCGAGATGAAGGCCCTTTCAGTGTGATGAATAAGCCAAAGGCTCTCTTCTGCTGAGGATG 1860  
DB 1801 TCCGAGATGAAGGCCCTTTCAGTGTGATGAATAAGCCAAAGGCTCTCTTCTGCTGAGGATG 1860

DB 1801 TCCGAGATGAAGGCTTCGATGATGAAGCCAAAGGTCCTCTCTCCCTCCGCTGAGATG 1860  
QY 1861 AAAACCTGAGGAAAAGGGAGCTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA 1920  
DB 1861 AAAACCTGAGGAAAAGGGAGCTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA 1920  
QY 1921 ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
DB 1921 ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
QY 1981 CAACAGGATGCAAGAGAGAGATCAAAATATTCATCATGACCCCGGAGCTCAGTGT 2040  
DB 1981 CAACAGGATGCAAGAGAGAGATCAAAATATTCATCATGACCCCGGAGCTCAGTGT 2040  
QY 2041 GGCCGACACAGGGCCCAAACTGCGAGCTCCGAATGACCTGGCTGGTGGATCCCA 2100  
DB 2041 GGCCGACACAGGGCCCAAACTGCGAGCTCCGAATGACCTGGCTGGTGGATCCCA 2100  
QY 2101 AGGAAGCTGCAAGATTCGATGTGCCAAGAGACAGGACCTGGAGGAAGCAAGGTGC 2160  
DB 2101 AGGAAGCTGCAAGATTCGATGTGCCAAGAGACAGGACCTGGAGGAAGCAAGGTGC 2160  
QY 2161 TCATCTTTGATGACTCCTTTTGAGCAGAGGTATGCGAGGATGCTCATCTTTCCGGCTGA 2220  
DB 2161 TCATCTTTGATGACTCCTTTTGAGCAGAGGTATGCGAGGATGCTCATCTTTCCGGCTGA 2220  
QY 2221 TATTCATGCTGGATGTGGCATCCGGAATGACACACAGCAGAGAGCGCTTCAG 2280  
DB 2221 TATTCATGCTGGATGTGGCATCCGGAATGACACACAGCAGAGAGCGCTTCAG 2280  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAA-CTCTGGAGAGA 2324  
DB 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAA-CTCTGGAGAGA 2324  
RESULT 3  
ID ADA00640 standard; cDNA; 2324 BP.  
AC ADA00640;  
XX  
DT 06-NOV-2003 (first entry)  
DE cDNA encoding human aspartyl (asparaginyl) beta-hydroxylase (HAAH).  
XX Tumour growth inhibition; human aspartyl (asparaginyl) beta-hydroxylase;  
KW HAAH hydroxylation; NOTCH polypeptide;  
KW epidermal growth factor-like repeat; EGF-like repeat; tumour cell;  
KW malignant neoplasm; colon cancer; breast cancer; pancreatic cancer;  
KW liver cancer; cancer of the bile duct; cancer the central nervous system;  
KW CNS; cytostatic; gene; ss; human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 12..2288  
FT /\*tag= a  
FT /\*product= "HAAH"  
XX  
PN US2003031670-A1.  
XX  
PD 13-FEB-2003.  
XX  
XX 08-NOV-1999; 99US-00436184.  
PF  
XX 08-NOV-1999; 99US-00436184.  
PR  
XX (WAND/) WANDS J R.  
PA (DMON/) DE LA MONTE S M.  
PA (INCE/) INCE N.  
PA (CARL/) CARLSON R I.  
XX

PI Wands JR, De La Monte SM, Ince N, Carlson RI;  
XX WPI; 2003-605701/57.  
DR P-PSDB; ADA00639.  
XX  
XX Inhibiting tumor growth or killing tumor cells (e.g. cancer of the colon,  
PT breast, pancreatic, liver or the central nervous system), by  
PT administering an inhibitor of the human aspartyl (asparaginyl) beta-  
PT hydroxylase.  
XX  
PS Disclosure; Page 3-4; 30pp; English.  
XX  
CC The present invention relates to a method for inhibiting tumour growth in  
CC a mammal. The method comprises administering to the mammal a compound,  
CC which inhibits the expression or enzymatic activity of a human aspartyl  
CC (asparaginyl) beta-hydroxylase (HAAH). The compound may inhibit HAAH  
CC hydroxylation of a NOTCH polypeptide. In particular, the compound may  
CC inhibit hydroxylation of an epidermal growth factor (EGF)-like repeat  
CC sequence in a NOTCH polypeptide. The methods are useful for inhibiting  
CC tumour growth or killing tumour cells, or for diagnosing or  
CC prognosticating a malignant neoplasm. In particular, the tumour or  
CC neoplasm is colon cancer, breast cancer, pancreatic cancer, liver cancer,  
CC cancer of the bile ducts, or cancer or tumour of the central nervous  
CC system (CNS). The present sequence encodes HAAH.  
XX  
SQ Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATATGGCCAGCGTAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60  
DB 1 CGGACCGTGCATATGGCCAGCGTAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60  
QY 61 GCTCCGGCAGCGGTAGCAGCGGTGCGGCGAGCAGCGCCCGGGCCCGGAGAGACAA 120  
DB 61 GCTCCGGCAGCGGTAGCAGCGGTGCGGCGAGCAGCGCCCGGGCCCGGAGAGACAA 120  
QY 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGACTCTCGGGAATCTTCTTCCACGT 180  
DB 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGACTCTCGGGAATCTTCTTCCACGT 180  
QY 181 GGTATTGCTGATGTCATTTGCTGGCGCTGAGACTCTGAGCTGCTGCTGCTGCTGCTG 240  
DB 181 GGTATTGCTGATGTCATTTGCTGGCGCTGAGACTCTGAGCTGCTGCTGCTGCTGCTG 240  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300  
DB 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300  
QY 301 ATTTTGTGATGATGTCACAAAGTTTATTAGGACTTTAAAGAGAGATCTTACTTCAGAGC 360  
DB 301 ATTTTGTGATGATGTCACAAAGTTTATTAGGACTTTAAAGAGAGATCTTACTTCAGAGC 360  
QY 361 CAGCAGTCCCGCCAGAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGCAGCTTCTGTG 420  
DB 361 CAGCAGTCCCGCCAGAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGCAGCTTCTGTG 420  
QY 421 AGGCAGAACCCGAGATATCGAAGTGAAGCAAGCAAGCAANTTCACTTCTCTCATG 480  
DB 421 AGGCAGAACCCGAGATATCGAAGTGAAGCAAGCAAGCAANTTCACTTCTCTCATG 480  
QY 481 AAATGGTACACGAGAACATGTTGAGGGAGAGAGCTTGCACAAAGAAAGATGGACCCAG 540  
DB 481 AAATGGTACACGAGAACATGTTGAGGGAGAGAGCTTGCACAAAGAAAGATGGACCCAG 540  
QY 541 GAGAACCAACAAAGAGAGATGATGATTTCTTATGGGAGCTGATGATGATGATGATG 600  
DB 541 GAGAACCAACAAAGAGAGATGATGATTTCTTATGGGAGCTGATGATGATGATGATG 600  
QY 601 AGACCTGGAACCTGAAGTATCTCATGAGAAACCGAGCATAGTTACCGCTGGAGAGA 660  
DB 601 AGACCTGGAACCTGAAGTATCTCATGAGAAACCGAGCATAGTTACCGCTGGAGAGA 660



```
Db 601 AGACCCGTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACACGTGGAAGAGA 660
Qy 661 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGTCCTGAGCAGGAAATCCAG 720
Db 661 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGTCCTGAGCAGGAAATCCAG 720
Qy 721 ATTCAGTGAACCAAGTAGTAGAAGATGAAGATGGAACCATGATACAGATGATTAACAT 780
Db 721 ATTCAGTGAACCAAGTAGTAGAAGATGAAGATGGAACCATGATACAGATGATTAACAT 780
Qy 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAAATCA 840
Db 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAAATCA 840
Qy 841 CAGAAGTAACCTGCTCCCTCAGGATATCTCTGAGAAATGTCACAGATTAATCTGAGAAG 900
Db 841 CAGAAGTAACCTGCTCCCTCAGGATATCTCTGAGAAATGTCACAGATTAATCTGAGAAG 900
Qy 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCACACAGAAACAAATAGAAAAA 960
Db 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCACACAGAAACAAATAGAAAAA 960
Qy 961 CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAAGAGCCTTAACCTTTTAATAAAT 1020
Db 961 CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAAGAGCCTTAACCTTTTAATAAAT 1020
Qy 1021 TTGATAAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCGTAAAAAGGGAAAAA 1080
Db 1021 TTGATAAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCGTAAAAAGGGAAAAA 1080
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGAAATACCCCTCAGAGTCCAGAG 1140
Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGAAATACCCCTCAGAGTCCAGAG 1140
Qy 1141 CAAGATATGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGTAATGAGG 1200
Db 1141 CAAGATATGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGTAATGAGG 1200
Qy 1201 TGCTACGTGGAGCCATCGAGACCTTACCAAGAGGTGGCCAGCCTACCTGATGCTCCCTGAG 1260
Db 1201 TGCTACGTGGAGCCATCGAGACCTTACCAAGAGGTGGCCAGCCTACCTGATGCTCCCTGAG 1260
Qy 1261 ACCTGCTGAAGCTGAGTTGAAGCGTGCCTGAGACAGGCAACAAATTTCTAGGTCAATAGA 1320
Db 1261 ACCTGCTGAAGCTGAGTTGAAGCGTGCCTGAGACAGGCAACAAATTTCTAGGTCAATAGA 1320
Qy 1321 GAGGTTCCCTGCTTACCTGAGAGATTTAGTCAACTATTTCCCAATGATACTTCCCTTAA 1380
Db 1321 GAGGTTCCCTGCTTACCTGAGAGATTTAGTCAACTATTTCCCAATGATACTTCCCTTAA 1380
Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATAATGACAATGCCAAAGAAAGTTT 1440
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATAATGACAATGCCAAAGAAAGTTT 1440
Qy 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATAGTGGCTTTGCTTAAAGTCCATTAATGGCTTCA 1500
Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATAGTGGCTTTGCTTAAAGTCCATTAATGGCTTCA 1500
Qy 1501 TCCTGAGGCAAGCAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAT 1560
Db 1501 TCCTGAGGCAAGCAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAT 1560
Qy 1561 CCGGAGATCTCTGGCACTGATGAGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620
Db 1561 CCGGAGATCTCTGGCACTGATGAGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620
Qy 1621 GGGTTGGGAAACAAAGAGGCATATAGTGGTATGAGCTTTGGGCACAAAGAGAGACACTTTG 1680
Db 1621 GGGTTGGGAAACAAAGAGGCATATAGTGGTATGAGCTTTGGGCACAAAGAGAGACACTTTG 1680
Qy 1681 CATCTGCTGCAACGCTCACTCTACAAATGTAATGGACTGAAAGCAGACAGCTTGGTGA 1740
Db 1681 CATCTGCTGCAACGCTCACTCTACAAATGTAATGGACTGAAAGCAGACAGCTTGGTGA 1740
```

```
Qy 1741 CCCCAAAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTGGAAGTTAA 1800
Db 1741 CCCCAAAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTGGAAGTTAA 1800
Qy 1801 TCCGAGATGAAGGCTTTGCAAGTATGATGATGAAGCAAGGCTCTTCTCCTCGCTGAGATG 1860
Db 1801 TCCGAGATGAAGGCTTTGCAAGTATGATGATGAAGCAAGGCTCTTCTCCTCGCTGAGATG 1860
Qy 1861 AAAACCTGAGGGAAGGAGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAAAGAA 1920
Db 1861 AAAACCTGAGGGAAGGAGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAAAGAA 1920
Qy 1921 ATGAAAAATGCTGCAAAAGGAGCTCTTAAAACTGTACCTTACTAGAAAAAGTTCCCGGAGA 1980
Db 1921 ATGAAAAATGCTGCAAAAGGAGCTCTTAAAACTGTACCTTACTAGAAAAAGTTCCCGGAGA 1980
Qy 1981 CAACAGAGATGAGAGGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040
Db 1981 CAACAGAGATGAGAGGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040
Qy 2041 GGCCGCAACACAGGCCCCACAAACTGCAAGGCTCCGAATGCACCTGGGCTTGGTGAATCCCA 2100
Db 2041 GGCCGCAACACAGGCCCCACAAACTGCAAGGCTCCGAATGCACCTGGGCTTGGTGAATCCCA 2100
Qy 2101 AGGAAGGCTCAAGATTCGATGTCGCAACAGAGCCAGGACCTGGGAGGAAGGAGTGC 2160
Db 2101 AGGAAGGCTCAAGATTCGATGTCGCAACAGAGCCAGGACCTGGGAGGAAGGAGTGC 2160
Qy 2161 TCATCTTTGATGACTCTTTTGACGACGAGGTATGGCAGGATGCCTCATCTTTCCGGCTGA 2220
Db 2161 TCATCTTTGATGACTCTTTTGACGACGAGGTATGGCAGGATGCCTCATCTTTCCGGCTGA 2220
Qy 2221 TATTATCGTGGAGTGTGGCATCCGGAACCTGACACACAGCAGAGACGAGCCTTCCAG 2280
Db 2221 TATTATCGTGGAGTGTGGCATCCGGAACCTGACACACAGCAGAGACGAGCCTTCCAG 2280
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324
```

## RESULT 4

ADP23914  
ID ADP23914 standard; cDNA; 2452 BP.

XX AC ADP23914;

XX DT 18-NOV-2004 (first entry)

XX DE PRO polypeptide encoding cDNA SEQ ID NO:1092.

XX KW ss: gene; PRO; antinflammatory; antiarthritic; antirheumatic;  
immunosuppressive; osteopathic; antidiabetic; dermatological;  
antiproliferative; antiallergic; antiasthmatic; hepatotropic; respiratory;  
gene therapy; immune system.

XX OS Unidentified.

XX PN WO2004041170-A2.

XX PD 21-MAY-2004.

XX PF 30-OCT-2003; 2003WO-US034312.

XX PR 01-NOV-2002; 2002US-0423394P.

XX PA (GETH ) GENENTECH INC.

XX PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
Wu TD;

XX XX WPI; 2004-419628/39.

DR P-PSDB; ADP23915.  
XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
PT renal disease, or demyelinating diseases of the central or peripheral  
PT nervous system.  
XX  
PS Claim 1; SEQ ID NO 1092; 2940pp; English.  
XX  
CC The invention relates to a novel isolated nucleic acid and the PRO  
CC polypeptide encoded by it. A protein of the invention has  
CC antiinflammatory, antirheumatic, antirheumatic, immunosuppressive,  
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
CC asthmatic, hepatotropic, and respiratory activity. A polynucleotide  
CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, athma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC Graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.  
XX  
SQ Sequence 2452 BP; 779 A; 507 C; 657 G; 509 T; 0 U; 0 Other;  
Query Match 95.6%; Score 2222; DB 13; Length 2452;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATGCGCCAGCGTAAGATGCCAGAGCAGCGGCAACAGCAGCAGCG 60  
DB 67 CGGACCGTGCATGCGCCAGCGTAAGATGCCAGAGCAGCGGCAACAGCAGCAGCG 126  
QY 61 GCTCCGCGCGGTAGCAGAGTGGCGGCGAGCAGCGCCCGCGGCGCGGAGAGACAA 120  
DB 127 GCTCCGCGCGGTAGCAGAGTGGCGGCGAGCAGCGCCCGCGGCGCGGAGAGACAA 186  
QY 121 AGCATGGAGGACCAAGAAATGGAGGAAAGCGGCACTCTCGGAACTTCAATCTTCACGT 180  
DB 187 AGCATGGAGGACCAAGAAATGGAGGAAAGCGGCACTCTCGGAACTTCAATCTTCACGT 246  
QY 181 GGTATTATGGTATGTCATGCTGGGCGTCTGGAATCTGTAGCTGTGTTGGTTGATC 240  
DB 247 GGTATTATGGTATGTCATGCTGGGCGTCTGGAATCTGTAGCTGTGTTGGTTGATC 306  
QY 241 TTGTTGACTATAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGTGTTGGTGGAG 300  
DB 307 TTGTTGACTATAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGTGTTGGTGGAG 366  
QY 301 ATTTTGATGGATGATGCCAAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
DB 367 ATTTTGATGGATGATGCCAAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGAGC 426  
QY 361 CAGAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTGTGG 420  
DB 427 CAGAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTGTGG 486  
QY 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAGAAAGAAATTCAGTCCCTTCATG 480

DB 487 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAGAAAGAAATTCAGTCCCTTCATG 546  
QY 481 AAATGGTACCGCAGAAACATGTTGAGGGAGAGACTTGGCAACAAGAGATGGACCCACAG 540  
DB 547 AAATGGTACCGCAGAAACATGTTGAGGGAGAGACTTGGCAACAAGAGATGGACCCACAG 606  
QY 541 GAGAACCAACCAAGAGAGATGATGTTCTTATGGCGACTGATGATGATGATGATGATG 600  
DB 607 GAGAACCAACCAAGAGAGATGATGTTCTTATGGCGACTGATGATGATGATGATGATG 666  
QY 601 AGACCTCTGGAACCTGGAAGTATCTCATGAAGAAACCGACATAGTTACCAGTGGAGAGA 660  
DB 667 AGACCTCTGGAACCTGGAAGTATCTCATGAAGAAACCGACATAGTTACCAGTGGAGAGA 726  
QY 661 CAGTTTTCACAGACTGTTAATCAGGATATGGAAGAGATGATGTTGAGCAGGAAATCCAG 720  
DB 727 CAGTTTTCACAGACTGTTAATCAGGATATGGAAGAGATGATGTTGAGCAGGAAATCCAG 786  
QY 721 ATTCCAGTGAACCAAGTATGGAAGATGGAAGATTGCCACCATGATACAGATGATGAACAT 780  
DB 787 ATTCCAGTGAACCAAGTATGGAAGATGGAAGATTGCCACCATGATACAGATGATGAACAT 846  
QY 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAATGAAGGATAGAATCA 840  
DB 847 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAATGAAGGATAGAATCA 906  
QY 841 CAGAAGTAACTGCTCCCTGAGGATTAATCTGTAGAGATTCACAGTAAATTTGTAAGAG 900  
DB 907 CAGAAGTAACTGCTCCCTGAGGATTAATCTGTAGAGATTCACAGTAAATTTGTAAGAG 966  
QY 901 AAGTAAAGATTTTCTGTGGAAGAACAGCAGAGAGATACCAACAGAAACAAATAGAAAAA 960  
DB 967 AAGTAAAGATTTTCTGTGGAAGAACAGCAGAGAGATACCAACAGAAACAAATAGAAAAA 1026  
QY 961 CAGATGATCCAGAACCAAAAAAGCAAGGTTAAGAAAAAGAGAGCCCTTAAATTTTAAAT 1020  
DB 1027 CAGATGATCCAGAACCAAAAAAGCAAGGTTAAGAAAAAGAGAGCCCTTAAATTTTAAAT 1086  
QY 1021 TTGATAAGACTATTAAAGCTGAACTTGTGCTGCAGAAAAAATCCGTAAAAAGGGAAAAA 1080  
DB 1087 TTGATAAGACTATTAAAGCTGAACTTGTGCTGCAGAAAAAATCCGTAAAAAGGGAAAAA 1146  
QY 1081 TTGAGGAAGCAGTGAATGATTAAGAACTAGTACCAATACCTCAGAGTCCACAG 1140  
DB 1147 TTGAGGAAGCAGTGAATGATTAAGAACTAGTACCAATACCTCAGAGTCCACAG 1206  
QY 1141 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
DB 1207 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1266  
QY 1201 TGCTACGTGAGCCATCGAGACCTACCAAGAGAGTGGCGAGCCTACCTGATGCTCCCTGCAG 1260  
DB 1267 TGCTACGTGAGCCATCGAGACCTACCAAGAGAGTGGCGAGCCTACCTGATGCTCCCTGCAG 1326  
QY 1261 ACCTGCTGAAGCTGAGTTGAAGCCTCCCTCAGACAGCAACAAATTTCTAGGTCAATGA 1320  
DB 1327 ACCTGCTGAAGCTGAGTTGAAGCCTCCCTCAGACAGCAACAAATTTCTAGGTCAATGA 1386  
QY 1321 GAGTTCCCTGCTTACCCTGCGAGATTAAGTAACTTATTTCCCAATGATATCTTCCTTAA 1380  
DB 1387 GAGTTCCCTGCTTACCCTGCGAGATTAAGTAACTTATTTCCCAATGATATCTTCCTTAA 1446  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATAATGACAAATGCAAGAAAGTTT 1440  
DB 1447 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATAATGACAAATGCAAGAAAGTTT 1506  
QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTGTAAAGTCCCAATTTATGGCTTCA 1500  
DB 1507 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTGTAAAGTCCCAATTTATGGCTTCA 1566  
QY 1501 TCCTGAGGCGACAGACAAATTTGCTGAGAGCATCCCATATTTTAAAGAGAGGATAGAT 1560  
DB 1567 TCCTGAGGCGACAGACAAATTTGCTGAGAGCATCCCATATTTTAAAGAGAGGATAGAT 1626

QY	1561	CCGAGATCTCTGGCACTGATGATCGGAGATTTATTTTCCACCTGGGGATGCCATGCGA	1620
Db	1627	CCGAGATCTCTGGCACTGATGATCGGAGATTTATTTTCCACCTGGGGATGCCATGCGA	1686
QY	1621	GGTTGGGAAACAAGAGGCATATAAGTGGTATGAGTTGGGCAACAAGAGGACACTTTG	1680
Db	1687	GGTTGGGAAACAAGAGGCATATAAGTGGTATGAGTTGGGCAACAAGAGGACACTTTG	1746
QY	1681	CATCTGTCTGCAACGCTCACTTACATGTGATGAGTGAAGCAAGCAAGCAAGCTTGGTGA	1740
Db	1747	CATCTGTCTGCAACGCTCACTTACATGTGATGAGTGAAGCAAGCAAGCAAGCTTGGTGA	1806
QY	1741	CCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAA	1800
Db	1807	CCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAA	1866
QY	1801	TCCAGATGAAGGCTTGCATGTATGATGATGAAGCAAGAGTCTCTCTGCTGAGATG	1860
Db	1867	TCCAGATGAAGGCTTGCATGTATGATGATGAAGCAAGAGTCTCTCTGCTGAGATG	1926
QY	1861	AAACCTGAGGAAAGGGGACTGGAGCCAGTTTCAGCTGTGCAGCAAGCAAGCAAGAA	1920
Db	1927	AAACCTGAGGAAAGGGGACTGGAGCCAGTTTCAGCTGTGCAGCAAGCAAGCAAGAA	1986
QY	1921	ATGAAATGCTGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA	1980
Db	1987	ATGAAATGCTGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA	2046
QY	1981	CAACAGGATGAGAAAGAGGACAGATCAAAATTTCCATCATGCAACCCGGGACTCACGTG	2040
Db	2047	CAACAGGATGAGAAAGAGGACAGATCAAAATTTCCATCATGCAACCCGGGACTCACGTG	2106
QY	2041	GGCCGACACAGGCCCCCAACTGCAGCTCCGAAATGCACCTGGGCTTGTGATTCCCA	2100
Db	2107	GGCCGACACACAGGCCCCCAACTGCAGCTCCGAAATGCACCTGGGCTTGTGATTCCCA	2166
QY	2101	AGGAAGGCTGCAAGATTCATGTGTCACACAGAGACCAAGCAAGCTGGGAGGAAGGCTGC	2160
Db	2167	AGGAAGGCTGCAAGATTCATGTGTCACACAGAGACCAAGCAAGCTGGGAGGAAGGCTGC	2226
QY	2161	TCATCTTTGATGACTCTTTGAGCAGAGGATATGACAGATGCCTCATCTTCCGGCTGA	2220
Db	2227	TCATCTTTGATGACTCTTTGAGCAGAGGATATGACAGATGCCTCATCTTCCGGCTGA	2286
QY	2221	TATTCATCTGTGATCTGTGGCATCCGAACTGACACCAAGAGAGAGAGAGAGAGAGAG	2280
Db	2287	TATTCATCTGTGATCTGTGGCATCCGAACTGACACCAAGAGAGAGAGAGAGAGAGAG	2346
QY	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTTCTGGAGAGA	2324
Db	2347	CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTTCTGGAGAGA	2390
RESULT 5			
ID	ADR97347		
XX	ADR97347 standard; DNA; 2452 BP.		
AC	ADR97347;		
XX			
DT	02-DEC-2004 (first entry)		
DE	Human ASPH DNA, an apoptosis related target Seq 55.		
XX			
KW	gene; ds; human; apoptosis; cancer; inflammation; autoimmune;		
KW	neurodegenerative disorder; cytostatic; antinflammatory;		
KW	immunopressive; neuroprotective; gene therapy; ASPH;		
XX	aspartate beta hydroxylase.		
OS	Homo sapiens.		
XX			
PN	WO2004078783-A2.		
XX			

PD	16-SEP-2004.		
XX			
PF	05-MAR-2004; 2004WO-GB000957.		
XX			
PR	07-MAR-2003; 2003GB-00005267.		
XX	(EIRX-) EIRX THERAPEUTICS LTD.		
PA			
XX	Murphy FU, Sheehan DE, Keating KE, Hayes I, Seera L;		
PI			
XX	WPI; 2004-662402/64.		
DR	P-PSDB; ADR97348.		
XX			
PT	Identifying an agent that modulates the function of an apoptosis-		
PT	associated polypeptide, useful for diagnosing or treating e.g. cancer,		
PT	comprises comparing the binding of the polypeptide to the candidate agent		
PT	and to a control agent.		
XX			
PS	Claim 2; SEQ ID NO 55; 304pp; English.		
XX			
CC	This invention relates to novel agents that modulates the function of		
CC	human apoptosis-associated proteins specified within the specification.		
CC	Specifically, it refers to a method for the identification of target		
CC	genes whose expression is correlated with an early stage in the		
CC	regulation of apoptosis. The present invention describes a method of		
CC	contacting either candidate agents or control agents to the target genes		
CC	and assessing the difference of binding and inhibitory activity, where		
CC	the candidate agent is selected from a low molecular weight organic		
CC	molecule, an antibody or fragment thereof, an antisense oligonucleotide,		
CC	a small inhibitory dsRNA, or a ribozyme. As such, the compositions and		
CC	methods are useful for diagnosing and treating diseases or conditions		
CC	associated with abnormal apoptosis in mammalian tissue, such as cancer,		
CC	inflammation, autoimmune or neurodegenerative disorders. Accordingly,		
CC	they exhibit cytostatic, antinflammatory, immunosuppressive and		
CC	neuroprotective activities. These may also be used for drug screening		
CC	purposes and in gene therapy. This polynucleotide sequence is a human		
CC	target gene associated with the regulation of apoptosis that can be		
CC	modulated by novel agents of the invention.		
XX			
SQ	Sequence 2452 BP; 779 A; 507 C; 657 G; 509 T; 0 U; 0 Other;		
	Query Match 95.6%; Score 2222; DB 13; Length 2452;		
	Best Local Similarity 99.9%; Pred. No. 0;		
	Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY	1	CGGACCGTGCATGCGCCAGCGTAAGAAATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG	60
Db	67	CGGACCGTGCATGCGCCAGCGTAAGAAATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG	126
QY	61	GCTCCGGCAGCGGTAGCACGAGTCGGGCGCAGCAGCAGCCCGGGCCCGGAGAGAGACAA	120
Db	127	GCTCCGGCAGCGGTAGCACGAGTCGGGCGCAGCAGCAGCCCGGGCCCGGAGAGAGACAA	186
QY	121	AGCATGGAGGACACAAGAATGGGAGGAAAGCGGACCTCTCGGGAACCTTCATTCTTCAGT	180
Db	187	AGCATGGAGGACACAAGAATGGGAGGAAAGCGGACCTCTCGGGAACCTTCATTCTTCAGT	246
QY	181	GGTTTATGGTGAATTCGATTCGTCGGCGCTCTGGCAATCTGTAGCTGTGCTTGGTTGATC	240
Db	247	GGTTTATGGTGAATTCGATTCGTCGGCGCTCTGGCAATCTGTAGCTGTGCTTGGTTGATC	306
QY	241	TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTATGGAG	300
Db	307	TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTATGGAG	366
QY	301	ATTTTGTATGTGGATGATGCCAAAGTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360
Db	367	ATTTTGTATGTGGATGATGCCAAAGTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC	426
QY	361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTAGCCCGCAGGAGCAGGTTCTCTGTTGG	420
Db	427	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTAGCCCGCAGGAGCAGGTTCTCTGTTGG	486

QY 421 AGGAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATG 480  
DB 487 AGGAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATG 546  
QY 481 AAATGGTACACGACAGAAACATGTTGAGGGAGAGAGCTTGCAACAAAGAGATGAGCCACAG 540  
DB 547 AAATGGTACACGACAGAAACATGTTGAGGGAGAGAGCTTGCAACAAAGAGATGAGCCACAG 606  
QY 541 GAGAACCAACAAAGAGAGATGATGATGTTCTTATGCGGACTGATGATGATGATGATGATG 600  
DB 607 GAGAACCAACAAAGAGAGATGATGATGTTCTTATGCGGACTGATGATGATGATGATGATG 666  
QY 601 AGACCTGGAACCTGAAGTATCTATGAGAAACCGAGCATGATGTTACCACTGGGAGAGAG 660  
DB 667 AGACCTGGAACCTGAAGTATCTATGAGAAACCGAGCATGATGTTACCACTGGGAGAGAG 726  
QY 661 CAGTTTCAACAAGCTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
DB 727 CAGTTTCAACAAGCTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 786  
QY 721 ATTCCAGTGAACCAAGTATGAGATCAAAAGATTGACCATGATGATGATGATGATGATGAT 780  
DB 787 ATTCCAGTGAACCAAGTATGAGATCAAAAGATTGACCATGATGATGATGATGATGATGAT 846  
QY 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTAGAAATGAGGGATGAGAAATCA 840  
DB 847 ACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTAGAAATGAGGGATGAGAAATCA 906  
QY 841 CAGAAGTAACTGCTCCCTGAGGATATCTGTGAGAGATTCACAGGATTAATTGTAGAG 900  
DB 907 CAGAAGTAACTGCTCCCTGAGGATATCTGTGAGAGATTCACAGGATTAATTGTAGAG 966  
QY 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGAGTACCAACAGAAACAAATAGAA 960  
DB 967 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGAGTACCAACAGAAACAAATAGAA 1026  
QY 961 CAGATGATCCAGAACAAAGAGTAAAGTAAAGAAAGAGCCTAACTTTAAATAAT 1020  
DB 1027 CAGATGATCCAGAACAAAGAGTAAAGTAAAGAAAGAGCCTAACTTTAAATAAT 1086  
QY 1021 TTGATAAGCACTTTAAAGCTGAACCTTGATGCTGCGAGAAATCTCCGTAAGGGGAAAA 1080  
DB 1087 TTGATAAGCACTTTAAAGCTGAACCTTGATGCTGCGAGAAATCTCCGTAAGGGGAAAA 1146  
QY 1081 TTGAGGAAGCACTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
DB 1147 TTGAGGAAGCACTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1206  
QY 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGAGAGTAAATGAGG 1200  
DB 1207 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGAGAGTAAATGAGG 1266  
QY 1201 TGCTACGTGGAGCCATCGACCTACCAAGGTGGCCAGCTACCTGATGCTCCCTGGAG 1260  
DB 1267 TGCTACGTGGAGCCATCGACCTACCAAGGTGGCCAGCTACCTGATGCTCCCTGGAG 1326  
QY 1261 ACCTGCTGAAGCTGAGTTGAAGCGTGGCTCAGACAGCAACAAATTTAGTGTCATATGA 1320  
DB 1327 ACCTGCTGAAGCTGAGTTGAAGCGTGGCTCAGACAGCAACAAATTTAGTGTCATATGA 1386  
QY 1321 GAGGTTCCCTGCTTACCTGACAGAGATGATTTCAACTATTTCCCAATGATCTCTCTTAA 1380  
DB 1387 GAGGTTCCCTGCTTACCTGACAGAGATGATTTCAACTATTTCCCAATGATCTCTCTTAA 1446  
QY 1381 AAAATGACCTTGGCTGGGATACCTCTTGATAGAGATGATGATGATGATGATGATGATG 1440  
DB 1447 AAAATGACCTTGGCTGGGATACCTCTTGATAGAGATGATGATGATGATGATGATGATG 1506  
QY 1441 ATGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500  
DB 1507 ATGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1566  
QY 1501 TCCTGAAGGACAGAAACAAATTTGCTGAGAGATCCCATATATTTAAAGAGAGATAGAT 1560

DB 1567 TCCTGAAGGACAGAAACAAATTTGCTGAGAGATCCCATATTTAAAGAGAGATAGAT 1626  
QY 1561 CCGAGATCTCGGCACTGATGATGAGGAGATTTTATTTTCACTGGGGATGCCATGAGA 1620  
DB 1627 CCGAGATCTCGGCACTGATGATGAGGAGATTTTATTTTCACTGGGGATGCCATGAGA 1686  
QY 1621 GGGTTGGAAACAAAGAGGCAATATAAGTGTGATGAGCTTGGGCAAGAGAGAGACACTTTG 1680  
DB 1687 GGGTTGGAAACAAAGAGGCAATATAAGTGTGATGAGCTTGGGCAAGAGAGAGACACTTTG 1746  
QY 1681 CATCTGTCTGGCAACGCTCACTCTACAAATGTGAATGAGCTTGAAGAGACAGCTTTGGTGA 1740  
DB 1747 CATCTGTCTGGCAACGCTCACTCTACAAATGTGAATGAGCTTGAAGAGACAGCTTTGGTGA 1806  
QY 1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGCTCTTTAGAAAGAACTGGAAGTTAA 1800  
DB 1807 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGCTCTTTAGAAAGAACTGGAAGTTAA 1866  
QY 1801 TCCGAGATGAAGGCTTTGCAAGTGAATGATGATGATGATGATGATGATGATGATGATG 1860  
DB 1867 TCCGAGATGAAGGCTTTGCAAGTGAATGATGATGATGATGATGATGATGATGATGATG 1926  
QY 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGGAAAGAA 1920  
DB 1927 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGGAAAGAA 1986  
QY 1921 ATGAAATGCTGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCTCCGAGA 1980  
DB 1987 ATGAAATGCTGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCTCCGAGA 2046  
QY 1981 CAAAGAGATGAGAGAGAGAGAGATCAAAATTTCCATGATGCAACCCCGGAGCTCAGTGT 2040  
DB 2047 CAAAGAGATGAGAGAGAGAGAGATCAAAATTTCCATGATGCAACCCCGGAGCTCAGTGT 2106  
QY 2041 GCGCGCACAGAGGCCCCCAAACTGAGGCTCCGAATGCACTGGGCTTGGTGTATTTCCCA 2100  
DB 2107 GCGCGCACAGAGGCCCCCAAACTGAGGCTCCGAATGCACTGGGCTTGGTGTATTTCCCA 2166  
QY 2101 AGGAAGCTGCAAGATTTCCATGTGCAACAGAGACAGAGCTGGAGAGAGGCAAGGTGC 2160  
DB 2167 AGGAAGCTGCAAGATTTCCATGTGCAACAGAGACAGAGCTGGAGAGAGGCAAGGTGC 2226  
QY 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGATGAGGATGCTCATCTTTCCGCTGA 2220  
DB 2227 TCATCTTTGATGATCTCTTTGAGCAGAGGATGATGAGGATGCTCATCTTTCCGCTGA 2286  
QY 2221 TATTCATGCTGGATGTGGGATCCGGAATCTGACACCAAGAGAGAGAGGCAAGCTTCCAG 2280  
DB 2287 TATTCATGCTGGATGTGGGATCCGGAATCTGACACCAAGAGAGAGAGGCAAGCTTCCAG 2346  
QY 2281 CAATTTAGCTGAATTTATGCAAGCTTGGGAACTCTGGAGAGA 2324  
DB 2347 CAATTTAGCTGAATTTATGCAAGCTTGGGAACTCTGGAGAGA 2390

## RESULT 6

ADJ56257

ID ADJ56257 standard; cDNA; 5358 BP.

XX AC ADJ56257;

XX AC ADJ56257;

DT 06-MAY-2004 (first entry)

XX DE Bovine cDNA differentially expressed in MYCN activated cells seqID 63.

XX KW bovine; differential expression; transactivator; proto-oncogene;

XX KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; sa;

XX OS MYCN activated cell.

XX OS Bos taurus.

PN US2003119009-A1.

XX PD 26-JUN-2003.  
XX PF 25-FEB-2002; 2002US-00084817.  
XX PR 23-FEB-2001; 2001US-0270784P.  
XX (STUA/) STUART S G.  
XX (NUCH/) NUCHTERN J G.  
XX (PLON/) PLON S E.  
XX (SHOH/) SHOHET J M.  
XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;  
XX WPI; 2003-635698/60.  
XX New genes regulated by MYCN activation, useful in gene therapy,  
XX particularly for treating a subject with e.g. neuroblastoma or other  
XX cancers, or for diagnosing, staging or monitoring the treatment of the  
XX cancer.  
XX Claim 1; SEQ ID NO 63; 27pp; English.  
XX This invention relates to novel isolated cDNAs that are differentially  
XX expressed in MYCN activated cells. Specifically, it refers to  
XX polynucleotide sequences that exhibit differential expression patterns in  
XX cells activated by the transactivator MYCN, where MYCN is a proto-  
XX oncogene that is amplified in neuroblastoma cells and is common in small  
XX cell lung cancers. The present invention describes these cDNA molecules  
XX as useful for in hybridisation assays to detect expression of nucleic  
XX acids (or complementary nucleic acids) in a present in a given sample, as  
XX well as for screening assays by identifying molecules or compounds that  
XX specifically bind the cDNA as a ligand and modulate function or activity.  
XX Accordingly, these compositions exhibit cytostatic activity and can also  
XX be used for gene therapy purposes. This polynucleotide sequence is a cDNA  
XX that is differentially expressed in MYCN activated cells, given in an  
XX exemplification of the invention. NOTE: This sequence does not appear in  
XX the Printed Specification but has been obtained in electronic format from  
XX the US Patent Office at  
XX ftp://seqdata.uspto.gov/sequence.html?docID=20030119009.  
XX SQ Sequence 5358 BP; 1684 A; 1009 C; 1204 G; 1461 T; 0 U; 0 Other;  
Query Match 93.4%; Score 2171; DB 10; Length 5358;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCAATGGCCCGAGCGTAAGAAATGCCAAGAGCAGCGCGCAACAGCAGCAGCG 60  
DB 258 CGGACCGTGCAATGGCCCGAGCGTAAGAAATGCCAAGAGCAGCGCGCAACAGCAGCAGCG 317  
QY 61 GCTCCGCGCAGCGTAGCAGAGTGGCGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 120  
DB 318 GCTCCGCGCAGCGTAGCAGAGTGGCGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 377  
QY 121 AGCATGGAGGACACAAGAAATGGAGAAAGGCGGACTCTCGGGAACCTTCATCTTCACGT 180  
DB 378 AGCATGGAGGACACAAGAAATGGAGAAAGGCGGACTCTCAGGAACTTCATCTTCACGT 437  
QY 181 GGTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
DB 438 GGTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 497  
QY 241 TTGTTGACTATGAGGAATGTTAGGAAATCTAGGAAATCTATGATGATGATGATGATGATGATGAT 300  
DB 498 TTGTTGACTATGAGGAATGTTAGGAAATCTAGGAAATCTATGATGATGATGATGATGATGATGAT 557  
QY 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
DB 558 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617  
QY 361 CAGCAGTCCCGCCAGAGAGCGGTAGGCCACACACTGAGCCCGGAGGAGCGAGGTTCTCTGTGG 420

DB 618 CAGCAGTCCCGCCAGAGAGCGGTAGGCCACACACTGAGCCCGAGGAGCGAGGTTCTCTGTGG 677  
QY 421 AGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
DB 678 AGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 737  
QY 481 AAATGGTACACGAGAAACATGTTGAGGGGAGAGACTTGCACAAAGAGATGAGCCACAG 540  
DB 738 AAATGGTACACGAGAAACATGTTGAGGGGAGAGACTTGCACAAAGAGATGAGCCACAG 797  
QY 541 GAGAACCAACAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 798 GAGAACCAACAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857  
QY 601 AGACCTCTGGAACCTGAAATGATCTCATGAAGAAACCGAGCATGATGATGATGATGATGATGAT 660  
DB 858 AGACCTCTGGAACCTGAAATGATCTCATGAAGAAACCGAGCATGATGATGATGATGATGATGAT 917  
QY 661 CAGTTTCAACAAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGATGAT 720  
DB 918 CAGTTTCAACAAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGATGAT 977  
QY 721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCACCATGATGATGATGATGATGATGAT 780  
DB 978 ATTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCACCATGATGATGATGATGATGATGAT 1037  
QY 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAATGAAGGATGAATCA 840  
DB 1038 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAATGAAGGATGAATCA 1097  
QY 841 CAGAGTAACTGCTCCCTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 1098 CAGAGTAACTGCTCCCTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1157  
QY 901 AAGTAAGCAATTTTCTCTGTGGAAGAAACAGCAGGAAGTACCACAGAAACAAATAGAAAA 960  
DB 1158 AAGTAAGCAATTTTCTCTGTGGAAGAAACAGCAGGAAGTACCACAGAAACAAATAGAAAA 1217  
QY 961 CAGATGATCCAGAAACAAAGCAAAAGTTAAGAAAGAGAGAGCTTAACTTTTAAATAAT 1020  
DB 1218 CAGATGATCCAGAAACAAAGCAAAAGTTAAGAAAGAGAGAGCTTAACTTTTAAATAAT 1277  
QY 1021 TTGATAAGACTATTAAAGCTGAACTTGATGCTGAGAAAAAATCCGTAAAGGGGAAAAA 1080  
DB 1278 TTGATAAGACTATTAAAGCTGAACTTGATGCTGAGAAAAAATCCGTAAAGGGGAAAAA 1337  
QY 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCCTCAGAGTCCACGAG 1140  
DB 1338 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCCTCAGAGTCCACGAG 1397  
QY 1141 CAAGATATGGAAGCGCGAGTGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
DB 1398 CAAGATATGGAAGCGCGAGTGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
QY 1201 TGCTACGTGGAGCCATCGAGACCTTACCAAGAGTGGCGAGCTACCTGATGCTCCTGAG 1260  
DB 1458 TGCTACGTGGAGCCATCGAGACCTTACCAAGAGTGGCGAGCTACCTGATGCTCCTGAG 1517  
QY 1261 ACCTGCTGAAAGCTGAGTTTGAAGCGTGTGAGGATGATGATGATGATGATGATGATGATGATGAT 1320  
DB 1518 ACCTGCTGAAAGCTGAGTTTGAAGCGTGTGAGGATGATGATGATGATGATGATGATGATGATGAT 1577  
QY 1321 GAGGTTCCCTGCTTACCTCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
DB 1578 GAGGTTCCCTGCTTACCTCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1637  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATGAGGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1638 AAAATGACCTTGGCGTGGGATACCTCTTGATGAGGATGATGATGATGATGATGATGATGATGAT 1697  
QY 1441 ATGAGAGAGTGTGAGTGTGACCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
DB 1698 ATGAGAGAGTGTGAGTGTGACCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1757

Qy	1501	TCCTGAAAGGCACAGAAACAAAATTGCTGAGAGCATCCCATATTTAAAGGAAGGATAGAAAT	1566
Db	1758	TCCTGAAGGCACAGAAACAAAATTGCTGAGAGCATCCCATATTTAAAGGAAGGATAGAAAT	1817
Qy	1561	CCGGAGATCTCGGCACTGATGCGGAGATTTATTTCCACCTGGGGGATGCCATCGAGA	1620
Db	1818	CCGGAGATCTCGGCACTGATGCGGAGATTTATTTCCACCTGGGGGATGCCATCGAGA	1877
Qy	1621	GGTTTGGGACAAAGAGGCAATATAGTGGTATGAGCTTGGGCACAAGAGAGACACTTTG	1680
Db	1878	GGTTTGGGACAAAGAGGCAATATAGTGGTATGAGCTTGGGCACAAGAGAGACACTTTG	1937
Qy	1681	CATCTGCTCTGGCAACGGCTCACTCTACAAATGGAATGGAATGAAAGCACAGGCTTGGTGG	1740
Db	1938	CATCTGCTCTGGCAACGGCTCACTCTACAAATGGAATGGAATGGAAGCACAGGCTTGGTGG	1997
Qy	1741	CCCAAAAGAAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
Db	1998	CCCAAAAGAAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	2057
Qy	1801	TCCGAGATCAAGGCCCTTCGAGTGATGGATTAAGCCGAAGGTCCTCTCTGCTGAGGATG	1860
Db	2058	TCCGAGATCAAGGCCCTTCGAGTGATGGATTAAGCCGAAGGTCCTCTCTGCTGAGGATG	2117
Qy	1861	AAAACCTGAGGGAAAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA	1920
Db	2118	AAAACCTGAGGGAAAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA	2177
Qy	1921	ATGAAAATGCTCTGCAAGAGGAGCTCTAAAAACCTGTACCTTACTAGAAAAGTTCCCGC	1980
Db	2178	ATGAAAATGCTCTGCAAGAGGAGCTCTAAAAACCTGTACCTTACTAGAAAAGTTCCCGC	2237
Qy	1981	CAACAGGATGCAGAGAGGACAGATCAAAATTTCCATCATGCAACCCGGGACTCACTGT	2040
Db	2238	CAACAGGATGCAGAGAGGACAGATCAAAATTTCCATCATGCAACCCGGGACTCACTGT	2297
Qy	2041	GGCCGACACAGGGGCCCAAACTGCAGGCTCGGAATGCACTGGGCTTGGTGAATCCCA	2100
Db	2298	GGCCGACACAGGGGCCCAAACTGCAGGCTCGGAATGCACTGGGCTTGGTGAATCCCA	2357
Qy	2101	AGGAAGGCTGCAAGATTGATGTGTCACAGCAACGAGACCTGGGAGGAAGCAAGGTGC	2160
Db	2358	AGGAAGGCTGCAAGATTGATGTGTCACAGCAACGAGACCTGGGAGGAAGCAAGGTGC	2417
Qy	2161	TCATCTTTGATGACTCTCTTTTGACACAGAGTATGGCAGGATGSCCTCATCTTCCGCTGA	2220
Db	2418	TCATCTTTGATGACTCTCTTTTGACACAGGATATGGCAGGATGSCCTCATCTTCCGCTGA	2477
Qy	2221	TATTCATCGTGATGCTGGGATCCGGACTGCACCAACGAGAGACGAGCGCTTCCAG	2280
Db	2478	TATTCATCGTGATGCTGGGATCCGGACTGCACCAACGAGAGACGAGCGCTTCCAG	2537
Qy	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA	2324
Db	2538	CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA	2581

## RESULT 7

RESULT 7  
ABX04178

ABX04178  
ID ABX04178 standard; cDNA: 2680 bp.

XX  
AC  
ABX04178;

DT 10-JAN-2003 (first entry)

XX DE Human mRNA differentially expressed in mesenchymal cells #25.

Human; ss; gene; skeletal growth; cartilage degeneration disorder;  
 chondroblastic phenotype; mesenchymal cell; cartilage formation;  
 bone formation; arthritis; osteoarthritis; rheumatoid arthritis;  
 gout arthritis; adjuvant arthritis; arthritis deformans; antigout;  
 infectious arthritis; osteochondrosis; RDA; antiarthritic; osteopathic;

KW	antirheumatic; antiinflammatory; representational difference analysis.
XX	
OS	Homo sapiens.
XX	
FN	WO200271927-A2.
XX	
PD	19-SEP-2002.
XX	
XX	
PF	12-MAR-2002; 2002WO-US007787.
XX	
PR	12-MAR-2001; 2001US-0274980P.
XX	
PA	(BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX	
PI	Yates KE, Mizuno S, Glowacki J;
XX	
DR	WPI; 2002-723276/78.
XX	
PT	New nucleic acid molecules capable of promoting chondrogenesis, useful
PT	for diagnosing and treating cartilaginous tissue degeneration conditions,
PT	e.g. osteoarthritis, rheumatoid arthritis, gout arthritis, or
PT	osterochondrosis.
XX	
XX	Claim 33; Page 123; 153pp; English.

Sequence 2680 BP; 854 A; 446 C; 575 G; 805 T; 0 U; 0 Other;

Query Match 38.5%; Score 895; DB 6; Length 2680;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 945; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY ' 1 CGGACCGTGCAATGGCCCAAGAGCGGCAACAGCAGCAGCG 60

Db  
1 CGGACCGTGCATGGCCACGGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG 60

61 GCTCCGGCAGCGGTAGCACGAGTCCGGCAGCAGCAGCCCGGGCCCGGACGACGACCA 130

[illegible]

131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

121











PR 13-DEC-2000; 2000US-00736457.  
 XX (BANG/) BANGUR C S.  
 PA (FANG/) FANGER G R.  
 PA (WANG/) WANG A.  
 PA (WANG/) WANG T.  
 PA (SWIT/) SWITZER A P.  
 PA (MCNEI/) MCNEILL P D.  
 PA (CLAP/) CLAPPER J D.  
 XX  
 PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;  
 PI Clapper JD;  
 XX  
 DR WPI; 2003-352750/33.  
 DR P-PSDB; ABU69516.  
 XX  
 PT Novel lung cancer polynucleotide encoding lung cancer protein, useful for  
 PT detecting the presence of lung cancer in a patient, and in pharmaceutical  
 PT compositions, e.g. vaccines, for treating lung cancer.  
 XX  
 PS Example 5; Page: 72pp; English.

The invention relates to a polynucleotide encoding a lung tumour protein, comprising a sequence selected from any of the 14 sequences mentioned in the specification, or a sequence (S2) mentioned in specification, complement of S1, sequences consisting of at least 20 contiguous residues of S1, sequences that hybridise to S1, sequences having 75% or preferably 90%, identity to S1, or degenerate variants of S1. Also included are an isolated polypeptide (comprising a sequence (S3) selected from any one of the 4 amino acid sequences mentioned in the specification, a sequence encoded by the polynucleotide, or sequences having at least 70%, preferably 90%, identity to a sequence encoded by the polynucleotide), an expression vector comprising the polynucleotide operably linked to an expression control sequence, a host cell transformed or transfected with the vector, an isolated antibody (or its antigen-binding fragment) that specifically binds to the polypeptide, detecting the presence of a cancer in a patient, a fusion protein comprising the polypeptide, an oligonucleotide that hybridises to S1 under moderately stringent conditions, stimulating and/or expanding T cells specific for a tumour protein (comprising contacting T cells with the polynucleotide, protein or antigen-presenting cells, under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells) and inhibiting the development of a cancer in a patient (by incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with the polynucleotide, protein or antigen presenting cells that express the polynucleotide, such that T cells proliferate, administering to the patient an effective amount of the proliferated T cells, and thus inhibiting the development of a cancer in the patient). The polynucleotide, protein and cells are useful in a composition for stimulating an immune response in a patient, and for treating a cancer in a patient (particularly lung cancer). The oligonucleotide is useful for determining the presence of a cancer in a patient. The protein and oligonucleotides are useful in pharmaceutical compositions, e.g. vaccines. The polynucleotide is also useful as a probe or primer for nucleic acid hybridisation, and in the design and preparation of ribozyme molecules for inhibiting expression of tumour polypeptides and proteins in tumour cells. An amplified portion of the polynucleotide is useful for isolating a full-length gene from a suitable library. The present sequence is a cDNA (full length, extended or partial) isolated from a library derived from lung tumour/cancer cells. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669

SQ Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;  
 Query Match 27 58; Score 640; DB 8; Length 2442;  
 Best Local Similarity 99.94; Pred. No. 6.5e-287;  
 Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 162 GGAACCTTCATCTTCACGCGTTTATGGTGATGTCATGCGCGCTGGACATCTGTA 221  
 Db 46 GGAACCTTCATCTTCACGCGTTTATGGTGATGTCATGCGCGCTGGACATCTGTA 105

QY 222 GCTGCGTTTGGTTTGATCTTGTTGACTATGAGGAGTTCTAGGAAACTAGGAATCTAT 281  
 Db 106 GCTGCGTTTGGTTTGATCTTGTTGACTATGAGGAGTTCTAGGAAACTAGGAATCTAT 165  
 QY 282 GATGCTGATGGTGATGAGAGATTTTGATGTGGATGATGCGAAAGTTTATTAGGACTTTAAA 341  
 Db 166 GATGCTGATGGTGATGAGAGATTTTGATGTGGATGATGCGAAAGTTTATTAGGACTTTAAA 225  
 QY 342 GAGAGATCTACTTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGGCACACACTGAGCCCC 401  
 Db 226 GAGAGATCTACTTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGGCACACACTGAGCCCC 285  
 QY 402 GAGGAGCAGGTTCTCTGTGGAGGAGAGAGCCAGAGATATCGAAGATGAGCAAGACACAA 461  
 Db 286 GAGGAGCAGGTTCTCTGTGGAGGAGAGAGCCAGAGATATCGAAGATGAGCAAGACACAA 345  
 QY 462 ATTCACTCCCTTCTCCATGAAATGTTACAGCGAGAGACATGTTGAGGAGAGAGACTTGCAG 521  
 Db 346 ATTCACTCCCTTCTCCATGAAATGTTACAGCGAGAGACATGTTGAGGAGAGAGACTTGCAG 405  
 QY 522 CAAGAAGATGAGCCACAGAGAGAACCAACAGAGAGATGATGAGTTTCTTATGGCGACT 581  
 Db 406 CAAGAAGATGAGCCACAGAGAGAACCAACAGAGAGATGATGAGTTTCTTATGGCGACT 465  
 QY 582 GATGTAGATGATAGATTGAGAGCCCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCAT 641  
 Db 466 GATGTAGATGATAGATTGAGAGCCCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCAT 525  
 QY 642 AGTTACCACTGGAGAGAGACAGTTTTCACAGAGCTGTTATCAGGATGAGGAGAGATGATG 701  
 Db 526 AGTTACCACTGGAGAGAGACAGTTTTCACAGAGCTGTTATCAGGATGAGGAGAGATGATG 585  
 QY 702 TCTGAGCAGGAAATCCAGATTCAGATGAAACAGTGTAGAGATGAGAGATGAGGACCAT 761  
 Db 586 TCTGAGCAGGAAATCCAGATTCAGATGAAACAGTGTAGAGATGAGAGATGAGGACCAT 645  
 QY 762 GATACAGATGATGTAAACATACCAAGTCTATGAGGAGAACAGCAGTATATGAACTCTAGAA 821  
 Db 646 GATACAGATGATGTAAACATACCAAGTCTATGAGGAGAACAGCAGTATATGAACTCTAGAA 705  
 QY 822 AATGAAGGATAGAAATTCACAGAGTAACTG 852  
 Db 706 AATGAAGGATAGAAATTCACAGAGTAACTG 736.

RESULT 12  
 ACA03258  
 ID ACA03258 standard; cDNA; 2442 BP.  
 XX AC ACA03258;  
 XX AC ACA03258;  
 DT 22-MAY-2003 (first entry)  
 DE Lung cancer therapy and diagnosis associated cDNA #1745.  
 XX DE Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.  
 XX OS Homo sapiens.  
 XX PN US2002172952-A1.  
 XX PD 21-NOV-2002.  
 XX PF 10-JUL-2001; 2001US-00902941.  
 XX 30-JUN-1999; 99US-00346492.  
 PR 15-OCT-1999; 99US-00419356.  
 PR 17-DEC-1999; 99US-00468867.  
 PR 30-DEC-1999; 99US-00476300.  
 PR 06-MAR-2000; 2000US-00519642.  
 PR 22-MAR-2000; 2000US-00533077.  
 PR 10-APR-2000; 2000US-00546259.  
 PR 27-APR-2000; 2000US-00560406.

```
PR 05-JUN-2000; 2000US-00589184.
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.
PR 26-SEP-2000; 2000US-006711325.
PR 06-OCT-2000; 2000US-00677419.
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX
XX (CORI-) CORIXA CORP.
XX
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;
XX Durham M, Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;
XX
XX WPI; 2003-328427/31.
XX
XX New polypeptide, useful for preparing a composition for treating or
XX inhibiting development of cancer, e.g. lung cancer.
XX
XX Example 5; SEQID NO 1791; 82pp; English.
XX
XX The invention describes an isolated polynucleotide comprising one of 32
XX sequences, complement or degenerate variants of them. The polynucleotide
XX is useful for preparing a composition e.g. a vaccine or for gene therapy,
XX for treating or inhibiting development of cancer, e.g. lung cancer. This
XX sequence represents a polynucleotide associated with the compositions and
XX methods for the therapy and diagnosis of lung cancer
XX
XX Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;
XX
XX Query Match 27.5%; Score 640; DB 8; Length 2442;
XX Best Local Similarity 99.9%; Pred. No. 6.5e-287;
XX Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 162 GGAACCTTCATCTTCACGTGGTTTATGGTGAATGCAATGCTGGCGCTGGACATCTGTA 221
DB 46 GGAACCTTCATCTTCACGTGGTTTATGGTGAATGCAATGCTGGCGCTGGACATCTGTA 105
QY 222 GCTGCTGTTGGTTTGAATCTTGTAATGATGAGGAAGTCTAGGAAACTAGGAATCTAT 281
DB 106 GCTGCTGTTGGTTTGAATCTTGTAATGATGAGGAAGTCTAGGAAACTAGGAATCTAT 165
QY 282 GATGCTGATGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 341
DB 166 GATGCTGATGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 225
QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACTGAGCCC 401
DB 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACTGAGCCC 285
QY 402 GAGGACAGGTTCCCTGTGAGGACAGAACCCAGAGATATCGAAGATGAAGCAAAAGACAA 461
DB 286 GAGGACAGGTTCCCTGTGAGGACAGAACCCAGAGATATCGAAGATGAAGCAAAAGACAA 345
QY 462 ATTCAGTCCCTTCATCAATATGGTACACGAGAACATGTTGAGGAGAACACTTGCAA 521
DB 346 ATTCAGTCCCTTCATCAATATGGTACACGAGAACATGTTGAGGAGAACACTTGCAA 405
QY 522 CAAGAAGATGGACCCACAGAGAACCAACAAAGAGGATGATGATGATGATGATGATGATGAT 581
DB 406 CAAGAAGATGGACCCACAGAGAACCAACAAAGAGGATGATGATGATGATGATGATGATGAT 465
QY 582 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641
DB 466 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525
QY 642 AGTTACACAGTGAAGAGACAGTTTCAAGACTGTTAATCAGATATGGAAGAGATGATG 701
DB 526 AGTTACACAGTGAAGAGACAGTTTCAAGACTGTTAATCAGATATGGAAGAGATGATG 585
QY 702 TCTGAGCAGGAAATCCAGATTCAGTGAACCACTAGTAGAGATCAAGAGATTCACCAT 761
DB 162 GGAACCTTCATCTTCACGTGGTTTATGGTGAATGCAATGCTGGCGCTGGACATCTGTA 221
DB 46 GGAACCTTCATCTTCACGTGGTTTATGGTGAATGCAATGCTGGCGCTGGACATCTGTA 105
QY 222 GCTGCTGTTGGTTTGAATCTTGTAATGATGAGGAAGTCTAGGAAACTAGGAATCTAT 281
```

```
DB 586 TCTGAGCAGGAAAAATCCAGATTCCAGTGAAACAGTAGTAGAAGATGAAGATTGCACCAT 645
QY 762 GATACAGATGATGTAACATACCAAGTCTATGAGGAACAAGCAGATGATATGAACCTCTAGAA 821
DB 646 GATACAGATGATGTAACATACCAAGTCTATGAGGAACAAGCAGATGATATGAACCTCTAGAA 705
QY 822 AATGAGGAGGATGAAATACAGAAAGTAACTG 852
DB 706 AATGAGGAGGATGAAATACAGAAAGTAACTG 736
XX
XX RESULT 13
XX ADH47310
XX ID ADH47310 standard; cDNA; 2442 BP.
XX AC ADH47310;
XX XX
XX DT 25-MAR-2004 (first entry)
XX XX Human lung tumour cDNA clone, SEQ ID NO 1791.
XX DE Human lung tumour cDNA clone, SEQ ID NO 1791.
XX XX lung cancer; tumour; cytostatic; immunostimulant; vaccine; gene therapy;
XX KW human; clone; ss.
XX KW Homo sapiens.
XX OS
XX XX
XX PN WO2003037267-A2.
XX XX
XX PD 08-MAY-2003.
XX XX
XX PF 28-OCT-2002; 2002WO-US034777.
XX XX
XX PR 29-OCT-2001; 2001US-00017754.
XX PR 28-MAR-2002; 2002US-00113872.
XX XX
XX XX (CORI-) CORIXA CORP.
XX PA
XX PI Henderson RA, Wang T, Watanabe Y, Kalos MD, Sleath PR;
XX Johnson JC, Retter MW, Durham M, Carter D, Fanger GR, Vedvick TS;
XX PI Bangur CS, McNabb A;
XX XX
XX DR WPI; 2003-468346/44.
XX XX
XX PT New polypeptides and encoding polynucleotides, useful for diagnosing,
XX preventing and/or treating lung cancer.
XX PT
XX XX Example 5; SEQ ID NO 1791; 258pp; English.
XX PS
XX XX The invention relates to novel compositions and methods for the therapy
XX and diagnosis of cancer, particularly lung cancer. The compositions
XX comprise one or more lung tumour polypeptides, immunogenic portions
XX thereof, polynucleotides that encode such polypeptides, antigen presenting
XX cells that express such polypeptides, and T cells that are specific for
XX cells expressing such polypeptides. The novel compositions have
XX cytostatic and immunostimulant activity. The lung tumour antigens can be
XX used in the creation of a vaccine. The polynucleotides that encode the
XX lung tumour polypeptides can be used in gene therapy to help in the
XX treatment of lung tumours. This polynucleotide sequence represents a
XX human lung tumour cDNA clone of the invention. This sequence was not
XX shown in the specification. It has been taken from a World Intellectual
XX Property Organization CD ROM supplied with the specification.
XX XX
XX XX Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;
XX SQ
XX
XX Query Match 27.5%; Score 640; DB 10; Length 2442;
XX Best Local Similarity 99.9%; Pred. No. 6.5e-287;
XX Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 162 GGAACCTTCATCTTCACGTGGTTTATGGTGAATGCAATGCTGGCGCTGGACATCTGTA 221
DB 46 GGAACCTTCATCTTCACGTGGTTTATGGTGAATGCAATGCTGGCGCTGGACATCTGTA 105
QY 222 GCTGCTGTTGGTTTGAATCTTGTAATGATGAGGAAGTCTAGGAAACTAGGAATCTAT 281
```



Qy	582	GATGTAGATGATAGATTTGAGACCTTGGAACTGAAGTATCTCATGAAGAAACCGAGCAT	641
Db	466	GATGTAGATGATAGATTTGAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT	525
Qy	642	AGTTACCACTGGGAAGACACAGTTCACAAAGCTGTATCAGGATATGGAAGAGATGATG	701
Db	526	AGTTACCACTGGGAAGACACAGTTCACAAAGCTGTATCAGGATATGGAAGAGATGATG	585
Qy	702	TCTCAGCAGGAAATCCAGATTCACAGTTCACAGTAGTAGAGATGAAAGATTGCACCAT	761
Db	586	TCTCAGCAGGAAATCCAGATTCACAGTTCACAGTAGTAGAGATGAAAGATTGCACCAT	645
Qy	762	GATACAGATGATGTAAACATACCAAGTCTATGAGGAACCAAGCAGTAGTATATGAACTCTAGAA	821
Db	646	GATACAGATGATGTAAACATACCAAGTCTATGAGGAACCAAGCAGTAGTATATGAACTCTAGAA	705
Qy	822	AATGAAGGGATAGAAATCACAGAAGTAACTG	852
Db	706	AATGAAGGGATAGAAATCACAGAAGTAACTG	736

## RESULT 15

ACN88788 standard: DNA: 3110 BP.

ACN88788:

DT 02-DEC-2004 (first entry)

DE Breast cancer related marker, seq id 9938.

XX  
Cancer: breast: tumour: cytostatic: marker: detection: therapy: ds.

XX Homo sapiens.

XX  
PN  
US2003099974-A1.

29-MAY-2003.

18-JUL-2002: 2002US-00198846.

18-JUL-2001: 2001US-0306220P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

WPI: 2003-787014/74.

Novel isolated polypeptide associated with breast cancer, useful for detecting presence of polypeptide in sample, as a marker for breast cancer.

PS Disclosure: SEO ID NO 9938; 36pp: English.

The invention relates to an isolated polypeptide (I) associated with breast cancer which is encoded by a nucleic acid molecule comprising a nucleotide sequence (S1). Further disclosed is an antibody that binds to the polypeptide of the invention. The activity of the polypeptide of the invention may be described as cytostatic. The antibody is useful for detecting the presence of (I) in a sample. Nucleic acid molecules of the invention are useful in the detection of breast tumours. (I) is useful as a marker for breast cancer and in breast cancer therapy. Sequences given in records ACN78951-ACN92934 represent nucleic acid markers associated with breast cancer. Note: The sequence listing does not form part of the specification but may be obtained in electronic format from the USPTO web site at [seqdata.uspto.gov/sequence.html?docID=20030093974](http://seqdata.uspto.gov/sequence.html?docID=20030093974)

Sequence 3110 BP; 1039 A; 488 C; 618 G; 915 T; 0 U; 50 Other;

Query Match      26.5%;    Score 615;    DB 11;    Length 3110;

Best Local Similarity 100.0%; Pred. No. 2.8e-275;

	Matches	615;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	332	AGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAAAGAGGCTCAGGCACA	391							
Db	595	AGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAAAGAGGCTCAGGCACA	654							
Qy	392	CACGTGACCCGAGGAGCAGGTTCCTGTGAGGCGAGAACCCAGAAATATCGAAGATGAAGC	451							
Db	655	CACGTGACCCGAGGAGCAGGTTCCTGTGAGGCGAGAACCCAGAAATATCGAAGATGAAGC	714							
Qy	452	AAAAGAAACAAATTCAGTCCCTTCCTCATGAAATGGTACAGCGAGAACATGTTTCAGGGAGA	511							
Db	715	AAAAGAAACAAATTCAGTCCCTTCCTCATGAAATGGTACAGCGAGAACATGTTTCAGGGAGA	774							
Qy	512	AGACTTGGCAACAAAGAGATGGACCCACAGAGAAACCAACAAGAGAGATGATCAGTTTCT	571							
Db	775	AGACTTGGCAACAAAGAGATGGACCCACAGAGAAACCAACAAGAGAGATGATCAGTTTCT	834							
Qy	572	TATGGCGCAGTGTAGATGATGATAGATTTGAGACCCCTGGAACTGGAAGTATCTCATGAAGA	631							
Db	835	TATGGCGCAGTGTAGATGATGATAGATTTGAGACCCCTGGAACTGGAAGTATCTCATGAAGA	894							
Qy	632	AACCGCAGCATAGTTTACCAGTGGAAAGAGACAGTTTCAAGACTGTAAATCAGGATATGGA	691							
Db	895	AACCGCAGCATAGTTTACCAGTGGAAAGAGACAGTTTCAAGACTGTAAATCAGGATATGGA	954							
Qy	692	AGAGATGATGTCGTGACGAGGAAATCCAGATTCAGTGGAACCACTAGTAGAAGATGAAAG	751							
Db	955	AGAGATGATGTCGTGACGAGGAAATCCAGATTCAGTGGAACCACTAGTAGAAGATGAAAG	1014							
Qy	752	ATTGCACCATGATACAGATGATGATTAACATACCAAGTCTATGAGGAAACAAGCAGTATATGA	811							
Db	1015	ATTGCACCATGATACAGATGATGATTAACATACCAAGTCTATGAGGAAACAAGCAGTATATGA	1074							
Qy	812	ACCTCTAGAAAAATGAAGGATAGAAATCACAAGATTAACGTCTCCCTCAGGATATATCC	871							
Db	1075	ACCTCTAGAAAAATGAAGGATAGAAATCACAAGATTAACGTCTCCCTCAGGATATATCC	1134							
Qy	872	TGTAGAAAGATTACAGGTAAATTCAGAAAGTAAGCATTTTTTCTGTGTGAAAGAACAGCA	931							
Db	1135	TGTAGAAAGATTACAGGTAAATTCAGAAAGTAAGCATTTTTTCTGTGTGAAAGAACAGCA	1194							
Qy	932	GGAAGTACCAACCAGA	946							
Db	1195	GGAAGTACCAACCAGA	1209							

## RESULT 16

RESOLUTION  
ACA92448  
ID ACA92448 standard: DNA: 2648 BP.

ACA92448:

DT 15-JUL-2003 (first entry)

DE DNA encoding human PMM-33.

Human; protein modification and maintenance molecule; PMW; cancer; cell proliferation disorder; atherosclerosis; neurological disorder; epilepsy; Huntington's disease; stroke; immune disorder; allergy; inflammatory disorder; AIDS; developmental disorder; hypothyroidism; Cushing's syndrome; gastrointestinal disorder; epithelial disorder; infection; cytostatic; antiarteriosclerotic; anticoagulant; nontropic neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnery; antiinflammatory; thyromimetic; gene; ds.

OS Homo sapiens.

PN WO2003031939-A2.

17-APR-2003.

11-OCT-2002; 2002WO-US032850.

12-OCT-2001; 2001US-0329689P.  
25-OCT-2001; 2001US-0335703P.  
09-NOV-2001; 2001US-0348887P.  
28-NOV-2001; 2001US-0334145P.  
06-DEC-2001; 2001US-0337451P.  
14-DEC-2001; 2001US-0340584P.  
(INCY-) INCYTE GENOMICS INC.  
Rankmar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;  
Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;  
Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
WPI; 2003-430274/40.  
P-PSDB; ABU92053.  
New human protein modification and maintenance molecules (PMMM), useful  
for diagnosing, treating and preventing diseases or conditions associated  
with the aberrant PMMM expression e.g. cancer, atherosclerosis, or  
infections.  
Claim 5; Page 306; 31pp; English.  
The present invention relates to the isolation of human protein  
modification and maintenance molecules (PMMM), and the polynucleotide  
sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM  
-1 to PMMM-40) are disclosed. The sequences of the invention are useful  
for diagnosing a condition or disease associated with the expression of  
PMMM in a subject, preparing a polyclonal or monoclonal antibody, and  
generating an expression profile of a sample containing the  
polynucleotides. The diseases or conditions associated with decreased  
expression or overexpression of PMMM are cell proliferation disorders  
(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,  
allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
syndrome), gastrointestinal or epithelial disorders, and infections. The  
PMMM polypeptides or their fragments are useful in screening compounds  
for effectiveness as agonists or antagonists of the polypeptides, or in  
altering the expression of the target polynucleotide and compounds that  
specifically bind to, or modulate the activity of the polypeptide.  
ACA92416-ACA92455 encode the human PMMM polypeptides of the invention  
Sequence 2648 BP; 844 A; 420 C; 545 G; 839 T; 0 U; 0 Other;  
Query Match 21.6%; Score 502; DB 10; Length 2648;  
Best Local Similarity 100.0%; Pred. No. 1e-222;  
Matches 502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
162 GGAACCTTCATTCTTACGTCGGTTTATGGTGATTGCGTGGCGCTCGACATCTGTA 221  
179 GGAACCTTCATTCTTACGTCGGTTTATGGTGATTGCGTGGCGCTCGACATCTGTA 238  
222 GCTGTCGTTTGGTTTGTATCTTTGTTGACTATAGGAAGTTCTAGCAAACTAGCAACTAT 281  
239 GCTGTCGTTTGGTTTGTATCTTTGTTGACTATAGGAAGTTCTAGCAAACTAGCAACTAT 298  
282 GATGTCGATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 341  
299 GATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 358  
342 GAGAGATCTACTTCAGAGCAGCAGTCGCCGCCAGAGAGGCTGAGCCACACTGAGCCC 401  
359 GAGAGATCTACTTCAGAGCAGCAGTCGCCGCCAGAGAGGCTGAGCCACACTGAGCCC 418  
402 GAGGAGCAGGTTCTCTGTCGAGGAGCAAGACCCAGCAATATCGAAGATGAAGCAAAAGACAA 461  
419 GAGGAGCAGGTTCTCTGTCGAGGAGCAAGACCCAGCAATATCGAAGATGAAGCAAAAGACAA 478  
462 ATTTCAGTCCCTCTTCCATGAATGTTACACGACAGCAACATGTTGAGGAGAGACTTGC 521

Db 479 ATTTCAGTCCCTCTTCCATGAATGTTACACGACAGCAACATGTTGAGGAGGAGAGACTTGC 538  
Qy 522 CAAGAAGATGACCCCAAGAGAGCAACCAACCAAGAGGATGATGATGATGATGATGATGATG 581  
Db 539 CAAGAAGATGACCCCAAGAGAGCAACCAACCAAGAGGATGATGATGATGATGATGATGATG 598  
Qy 582 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641  
Db 599 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658  
Qy 642 AGTTACACGTCGGAGAGACAG 663  
Db 659 AGTTACACGTCGGAGAGACAG 680  
RESULT 17  
ACH47067  
ID ACH47067 standard; cDNA; 502 BP.  
XX  
AC ACH47067;  
XX  
DT 13-OCT-2003 (first entry)  
XX  
DE Human infant brain cDNA #1130.  
XX  
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
genome mapping; biodiversity; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2003073623-A1.  
XX  
PD 17-APR-2003.  
XX  
PF 30-JUL-2001; 2001US-00918995.  
XX  
PR 30-JUL-2001; 2001US-00918995.  
XX  
PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
DR WPI; 2003-615964/58.  
XX  
PT New polynucleotide sequences obtained from various cDNA libraries, useful  
as hybridization probes, as oligomers for PCR, for chromosome and gene  
mapping, in the recombinant production of protein, or in generating  
antisense DNA or RNA.  
XX  
PS Claim 1; SEQ ID NO 34279; 44pp; English.  
XX  
CC The invention relates to an isolated polynucleotide comprising any one of  
38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
determined by the technique of SBH (sequencing by hybridisation). Also  
included is a purified polypeptide comprising a sequence corresponding to  
a reading frame of the novel polynucleotide. The nucleic acid sequences  
are useful in diagnostics as expressed sequence tags (EST) for  
identifying expressed genes or for physical mapping of the human genome,  
in forensics, in assessing biodiversity, or in identifying mutations  
responsible for genetic disorders and other traits. The nucleotide  
sequences are also useful as hybridisation probes, as oligomers for PCR,  
for chromosome and gene mapping, in the recombinant production of  
protein, or in generating antisense DNA or RNA. The purified polypeptide  
is useful for generating antibodies specific for it. The present sequence  
is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
for this patent did not form part of the printed specification, but was  
obtained in electronic format directly from USPTO at  
seqdata.uspto.gov/sequence.html?DocID=20030073623

SQ	Sequence 502 BP; 178 A; 97 C; 119 G; 106 T; 0 U; 2 Other;	
	Query Match 19.8%; Score 459; DB 9; Length 502;	
	Best Local Similarity 100.0%; Pred. No. 1e-202;	
	Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	912 TTTCCTGTGGGAAGAACACGACGGAAGTACACACAGAAACAAATAGAAAAACAGATGATCCA	971
Db	44 TTTCCTGTGGGAAGAACACGACGGAAGTACACACAGAAACAAATAGAAAAACAGATGATCCA	103
Qy	972 GAACAAAAGCAAAAGTTTAAAGAAAAGAGCCTAAACTTTTAAATAAATTTGATAAGACT	1031
Db	104 GAACAAAAGCAAAAGTTTAAAGAAAAGAGCCTAAACTTTTAAATAAATTTGATAAGACT	163
Qy	1032 ATTAAGCTGAACCTTGATCTGCAGAAAACCTCCGTAAGGGGGAAGAAATTTGAGGAGCA	1091
Db	164 ATTAAGCTGAACCTTGATCTGCAGAAAACCTCCGTAAGGGGGAAGAAATTTGAGGAGCA	223
Qy	1092 GTGAATGCATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAGCAAGATATGGG	1151
Db	224 GTGAATGCATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAGCAAGATATGGG	283
Qy	1152 AAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATATAGAGTGTCTACGTGGA	1211
Db	284 AAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATATAGAGTGTCTACGTGGA	343
Qy	1212 GCCATCGAGACCTACCAAGAGTGGCCAGCCTACTGTGATGTCCTGCGAGACCTGCTGAAG	1271
Db	344 GCCATCGAGACCTACCAAGAGTGGCCAGCCTACTGTGATGTCCTGCGAGACCTGCTGAAG	403
Qy	1272 CTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGTTCATATGAGAGGTTCCCTG	1331
Db	404 CTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGTTCATATGAGAGGTTCCCTG	463
Qy	1332 CTTACCTGCGAGAGATTAGTTCAACTATTTCCTCCCAATGAT 1370	
Db	464 CTTACCTGCGAGAGATTAGTTCAACTATTTCCTCCCAATGAT 502	
RESULT 18		
ABQ59507/c		
ID	ABQ59507 standard; cDNA; 660 BP.	
AC	ABQ59507;	
XX	02-AUG-2002 (first entry)	
DT	Human colon cancer related nucleotide sequence SEQ ID NO:3202.	
DE	Human; colon cancer; cancer; tissue profiling; forensic; mapping;	
KW	genetic analysis; diagnostic; antisense therapy; gene; ss.	
XX	Homo sapiens.	
OS	WO200229086-A2.	
PN	11-APR-2002.	
PD	02-OCT-2001; 2001WO-US030732.	
PF	02-OCT-2000; 2000US-0237271P.	
XX	(FARB ) BAYER CORP.	
PA	Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;	
XX	Thiagalingam A, Lewis ME;	
PI	WPI; 2002-426115/45.	
XX	New isolated nucleic acid that is differentially expressed in cancer	
PT	tissues useful for determining the presence of colon cancer in a cell or	
PT	tissue type, and in antisense therapy.	
XX		
PS	Claim 1; Fig 1; 796pp; English.	
XX	ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially	
CC	expressed in cancer tissues. AB578993 to AB579004 represent proteins	
CC	encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be	
CC	used in antisense therapy. An antibody immunoreactive with a polypeptide	
CC	encoded by (I) is useful for detecting cancer in a patient sample, and	
CC	for detecting the presence or absence of a polynucleotide encoded by a	
CC	nucleic acid which hybridises to (I) in a cell. A probe/primer derived	
CC	from (I) can be used for determining the presence of a nucleic acid which	
CC	hybridises to (I), and for determining the phenotype of cells in a sample	
CC	of cells from a patient. (I) is useful for determining the presence of	
CC	colon cancer in a cell or tissue type, for determining the presence or	
CC	state of other type of cancer, in antisense therapy, to generate	
CC	macroarrays on a solid surface, to identify a chromosome on which the	
CC	corresponding gene resides, and in tissue profiling, forensics, genetic	
CC	analysis, mapping and diagnostic applications. (I) can be used to raise	
CC	antibodies, and to screen for peptide analogues and antagonists	
XX		
SQ	Sequence 660 BP; 131 A; 181 C; 123 G; 221 T; 0 U; 4 Other;	
	Query Match 15.4%; Score 359; DB 6; Length 660;	
	Best Local Similarity 100.0%; Pred. No. 3.4e-156;	
	Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	581 TGATGTAGATGATAGATTTTCAGAGACCCCTGGAAACCTGGAATATCTCATGAAGAACCCGACCA	640
Db	462 TGATGTAGATGATAGATTTTCAGAGACCCCTGGAAACCTGGAATATCTCATGAAGAACCCGACCA	403
Qy	641 TAGTTACACGTGGAAGAGACAGTTTCAAAAGACTGTAAATCAGGATATGGAAGAGATGAT	700
Db	402 TAGTTACACGTGGAAGAGACAGTTTCAAAAGACTGTAAATCAGGATATGGAAGAGATGAT	343
Qy	701 GTCTGACGAGAAATCCAGATTCAGTGAACCAAGTATAGAGAAACAAGAGATTGACCA	760
Db	342 GTCTGACGAGAAATCCAGATTCAGTGAACCAAGTATAGAGAAACAAGAGATTGACCA	283
Qy	761 TGATACAGATGATGTAAACATACCAAGTCTATGAGGAAACAAGAGATTATGAACCTCTAGA	820
Db	282 TGATACAGATGATGTAAACATACCAAGTCTATGAGGAAACAAGAGATTATGAACCTCTAGA	223
Qy	821 AAATGAAGGGATGAAATATCAGAGAGTAACTGCTCCCTCGAGGATATCTCTGAGAAGA	880
Db	222 AAATGAAGGGATGAAATATCAGAGAGTAACTGCTCCCTCGAGGATATCTCTGAGAAGA	163
Qy	881 TTCACAGTAAATGTAGAAGAAAGTAAGCATTTTCTGTGTGGAAGAACAGCAGGAGATAC	939
Db	162 TTCACAGTAAATGTAGAAGAAAGTAAGCATTTTCTGTGTGGAAGAACAGCAGGAGATAC	104
RESULT 19		
ADT94758		
ID	ADT94758 standard; cDNA; 495 BP.	
XX	AC ADT94758;	
XX	16-DEC-2004 (first entry)	
DT	Colon cancer associated human cDNA sequence #277.	
DE	Colon cancer; T cell; tumour protein; C634S; C637S; C640S; C636S;	
XX	humoral immune response; cellular immune response; cytostatic;	
KW	immunostimulant; human; ss.	
XX	Homo sapiens.	
OS	US2003087818-A1.	
PN	08-MAY-2003.	
PD	01-FEB-2002; 2002US-00066543.	
XX	02-FEB-2001; 2001US-0267400P.	
PR		





cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at [www.uspto.gov/sequence.html?DocID=20030194704](http://www.uspto.gov/sequence.html?DocID=20030194704)

XX	
SQ	Sequence 592 BP; 220 A; 122 C; 79 G; 171 T; 0 U; 0 Other; Query Match           4.5%; Score 104; DB 12; Length 592; Best Local Similarity 100.08; Pred. No. 1.4e-37; Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	162 GGAACTTCATTCTTCAAGTGTTATCGGTGATTCGATTGGCGCCTCGGCACATCTGTA 221
Db	230 GGAACTTCATTCTTCAAGTGTTATCGGTGATTCGATTGGCGCCTCGGCACATCTGTA 171
QY	222 GCCTGTCGTTTTGGTTTGACTTTGTTGACTATGAGAAGTCTTAGG 265
bB	170 GCTGTCGTTTTGGTTTGACTTTGTTGACTATGAGAAGTCTTAGG 127

PD	14-JAN-1999.	
XX		
XX		
PF	30-JUN-1998;	98WO-US013608.
XX		
XX	01-JUL-1997;	97US-0051381P.
PR	01-JUL-1997;	97US-0051480P.
PR	12-SEP-1997;	97US-0058598P.
PR	12-SEP-1997;	97US-0058663P.
XX		
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA		
XX		
XX	Carter KC, Feng P, Rosen CA, Ruben SM, Endress GA;	
DI		
XX		
XX	WPI; 1999-105683/09.	
DR	P-PSDB; AAY01148, AAY01190, AAY01191, AAY01192, AAY01193.	
DR		
XX		
PT	New isolated human genes and the secreted polypeptides they encode -	
PT	useful for diagnosis and treatment of e.g. cancers, neurological	
PT	disorders, immune diseases, immune deficiency diseases or blood	
PT	disorders.	
PT		
XX		
XX	Claim 4; Page 131; 179pp; English.	
PS		

Claim 4; Page 131; 179pp; English.

The invention relates to nucleic acid sequences (AAx22111 to AAx22134) encoding human secreted proteins (AAy01135 to AAy01158). The secreted protein gene sequences are deposited with the ATCC under deposit number ATCC 209118. Host cells comprising recombinant vectors containing the nucleic acid sequences are used for the recombinant production of the secreted proteins. The polynucleotide and amino acid sequences are useful for are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Pathological conditions can be also diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the polynucleotides. Specific uses are described for each of the polynucleotides, based on which tissues they are most highly expressed in, and include developing products for the diagnosis or treatment of cancer, tumours, developmental abnormalities and foetal deficiencies, autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders, schizophrenia, immunological disorders, immune deficiency diseases, (AIDS), mood disorders, respiratory disorders, arthritis, asthma, haematopoietic disorders, neural disorders, skeletal disorders, osteoporosis, metabolic disorders, cardiovascular disorders, endocrine disorders or gastrointestinal disorders. The polypeptides are also useful for identifying their binding partners. The present sequence represents a gene encoding a human secreted protein (see descriptor line for gene number and clone identification)



Db 293 AGTTTATTAGG 304

RESULT 26  
ADN95564/c  
ID ADN95564 standard; DNA; 60 BP.  
XX  
AC ADN95564;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Human BEC/LEC-related gene sequence SeqID487.  
XX  
KW growth; differentiation; blood endothelial cell; BEC;  
KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;  
KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;  
KW vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;  
KW inflammatory disease; cancer metastasis; lymphatic system; gene; ds;  
KW human.  
XX  
OS Homo sapiens.  
XX  
PN WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 20-JUL-2001; 2001WO-IB001903.  
XX  
PR 28-JUL-2000; 2000US-0221607P.  
XX  
PR 02-MAY-2001; 2001US-0287724P.  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX  
PI WPI; 2002-257383/30.  
XX  
DR  
XX  
PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of a  
PT genome, useful for detecting tissue-, pathology-, and developmental-  
PT specific genes.  
XX  
PS Example 1; SEQ ID NO 11070; 47pp; English.  
XX  
CC The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-  
CC )transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises several  
CC oligonucleotides, each capable of hybridising selectively to a set of  
CC messenger RNAs transcribed from a given transcription unit of the genome,  
CC which encodes one or more messenger RNA splice variants. The  
CC oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterising the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a particular  
CC - and pathology-specific genes, such as those genes only expressed in  
CC specific tissue under a specific pathological condition; to detect  
CC developmental specific genes; and to detect RNA transcripts and splice  
CC variants of a transcriptome of a patient suffering from a particular  
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from  
CC rats, humans and mice, which are used in the exemplification of the  
CC present invention. N.B. The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIFO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 60 BP; 22 A; 8 C; 17 G; 13 T; 0 U; 0 Other;  
Query Match 2.6%; Score 60; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 4.2e-17;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 672 GACTTATCAGGATATGGAAGAGATGATCTGACGAGGAATTCAGATCCAGTGA 731  
Db 1 GACTTATCAGGATATGGAAGAGATGATCTGACGAGGAATTCAGATCCAGTGA 60

RESULT 27  
ADN95564/c  
ID ADN95564 standard; DNA; 2208 BP.  
XX  
AC ADN95564;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Human BEC/LEC-related gene sequence SeqID487.  
XX  
KW growth; differentiation; blood endothelial cell; BEC;  
KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;  
KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;  
KW vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;  
KW inflammatory disease; cancer metastasis; lymphatic system; gene; ds;  
KW human.  
XX  
OS Homo sapiens.  
XX  
PN WO2003080640-A1.  
XX  
PD 02-OCT-2003.  
XX  
PF 07-MAR-2003; 2003WO-US006900.  
XX  
PR 07-MAR-2002; 2002US-0363019P.  
XX  
PI (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI (LICN) LICENTIA LTD.  
XX  
PI Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;  
XX  
PI WPI; 2003-876899/81.  
XX  
DR P-PSDB; ADN95563.  
XX  
PS Example 1; SEQ ID NO 487; 176pp; English.  
XX  
CC This invention relates to a method of differentially modulating the  
CC growth or differentiation of blood endothelial cells (BEC) or lymphatic  
CC endothelial cells (LEC) comprising contacting endothelial cells with a  
CC composition comprising an agent that differentially modulates blood or  
CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises  
CC identifying a human subject with lymphoedema and with a mutation in at  
CC least one allele of a gene encoding a LEC protein, where the proviso that  
CC correlates with lymphoedema in human subjects, and with the proviso that  
CC the LEC protein is not VEGFR-3; and administering to the subject a  
CC composition comprising a lymphatic growth agent selected from VEGF-C or  
CC VEGF-D polypeptides and polynucleotides. The invention may be useful for  
CC the development of compounds with an antiangiogenic, cytostatic,  
CC vasotropic or antiinflammatory activity or for gene therapy. The method  
CC is useful in modulating the growth or differentiation of blood  
CC endothelial cells or lymphatic endothelial cells; in treating hereditary  
CC lymphoedema, in screening for an endothelial cell disorder or  
CC predisposition to the disorder or in monitoring the efficacy or toxicity  
CC of a drug on endothelial cells. The agent is useful in manufacturing a  
CC medicament for the differential modulation of blood vessel endothelial  
CC cell or lymphatic vessel endothelial cell growth or differentiation. The  
CC lymphatic growth agent may also be used in manufacturing a medicament for  
CC the treatment of hereditary lymphoedema resulting from a mutation in a  
CC LEC gene or of other diseases involving the lymphatic vessels, such as  
CC various inflammatory diseases and cancer metastasis via the lymphatic  
CC system. The present sequence is that of a human LEC/BEC differentially  
CC expressed gene which is related to the method of the invention. Note: This  
CC sequence does not appear in the specification but was obtained by the  
CC indexer using the source data given in table 14 of the specification.  
XX  
SQ Sequence 2208 BP; 621 A; 428 C; 484 G; 675 T; 0 U; 0 Other;

Query Match 2.5%; Score 59; DB 11; Length 2208;  
Best Local Similarity 100.0%; Pred. No. 1.1e-16;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC	survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC	bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC	syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC	immune activity, disorders related to aberrant acute phase responses,
CC	hypergenital conditions, birth defects, necrotic lesions, wounds,
CC	organ transplant rejection, conditions related to organ transplant
CC	rejection, disorders related to aberrant signal transduction,
CC	proliferating disorders, cancers and HIV propagation in cells infected
CC	with other viruses. The present sequence is that of a human gene which is
CC	subject to the novel association with the NF-kappaB pathway of the
CC	invention. Note: This sequence does not appear in the specification but
CC	was obtained by the indexer from Genbank.
XX	
SQ	Sequence 2208 BP; 621 A; 428 C; 484 G; 675 T; 0 U; 0 Other;
	Query Match 2.5%; Score 59; DB 13; Length 2208;
	Best Local Similarity 100.0%; Pred. No. 1.1e-16;
	Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	2239 GGCATCCGGAACTGCACACACAGACAGACGCGACGCTTCACGCAATTAGCATGAATTC 2297
Db	59 GGCATCCGGAACTGCACACACAGACAGACGCGACGCTTCACGCAATTAGCATGAATTC 1
RESULT 29	
ADP23153/c	
ID	ADP23153 standard; cDNA; 2208 BP.
XX	
AC	ADP23153;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	PRO polypeptide encoding cDNA SEQ ID NO:247.
XX	
KW	ss; gene; PRO: antiinflammatory; antiarthritic; antirheumatic;
KW	immunosuppressive; osteopathic; antidiabetic; dermatological;
KW	antipsoriatic; anciallergic; antiasthmatic; hepatotropic; respiratory;
KW	gene therapy; immune system.
XX	
OS	Unidentified.
XX	
PN	W02004041170-A2.
XX	
PD	21-MAY-2004.
XX	
PF	30-OCT-2003; 2003WO-US034312.
XX	
PR	01-NOV-2002; 2002US-0423394P.
XX	
PA	(GETH ) GENENTECH INC.
XX	
PI	Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI	Wu TD;
XX	
DR	WPI: 2004-419628/39.
DR	P-PSDB; ADP23154.
XX	
PT	New PRO polypeptides and polynucleotides, useful for treating e.g.
PT	erythematosis, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT	renal disease, or demyelinating diseases of the central or peripheral
PT	nervous system.
XX	
PS	Claim 1; SEQ ID NO 247; 2940pp; English.
XX	
CC	The invention relates to a novel isolated nucleic acid and the PRO
CC	polypeptide encoded by it. A protein of the invention has
CC	antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC	osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC	antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC	of the invention may have a use in gene therapy. The PRO polypeptide, its
CC	agonist, antagonist, or antibody that specifically binds to the
CC	polypeptide is useful for treating an immune related disorder such as
CC	systemic lupus erythematosis, rheumatoid arthritis, osteoarthritis,

CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.

XX Sequence 2208 BP; 621 A; 428 C; 484 G; 675 T; 0 U; 0 Other;

Query Match 2.5%; Score 59; DB 13; Length 2208;  
Best Local Similarity 100.0%; Pred. No. 1.1e-16;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2239 GGCATCCGCACTGACACACAGAGAGCGAGCGCTTCCAGCAATTTAGCATGATTC 2297  
DB 59 GGCATCCGCACTGACACACAGAGAGCGAGCGCTTCCAGCAATTTAGCATGATTC 1

RESULT 30  
ABQ59571/c  
ID ABQ59571 standard; cDNA; 583 BP.  
XX AC ABQ59571;  
XX 02-AUG-2002 (first entry)  
XX Human colon cancer related nucleotide sequence SEQ ID NO:3266.  
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
XX genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX Homo sapiens.  
XX WO200229086-A2.  
XX 11-APR-2002.  
XX 02-OCT-2001; 2001WO-US030732.  
XX 02-OCT-2000; 2000US-0237271P.  
XX (FARB ) BAYER CORP.  
XX Burgess C, Aetle JH, Carroll E, Catino TU, Dwivedi P, Molino GA;  
XX Thiagalingam A, Lewis ME;  
XX WPI; 2002-426115/45.

XX New isolated nucleic acid that is differentially expressed in cancer  
XX tissues useful for determining the presence of colon cancer in a cell or  
XX tissue type, and in antisense therapy.  
XX Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
XX expressed in cancer tissues. AB578993 to AB579004 represent proteins  
XX encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
XX used in antisense therapy. An antibody immunoreactive with a polypeptide  
XX encoded by (I) is useful for detecting cancer in a patient sample, and  
XX for detecting the presence or absence of a polynucleotide encoded by a  
XX nucleic acid which hybridises to (I) in a cell. A probe/primer derived

CC from (I) can be used for determining the presence of a nucleic acid which  
CC hybridises to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence of  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists

XX Sequence 583 BP; 170 A; 128 C; 126 G; 152 T; 0 U; 7 Other;

Query Match 2.2%; Score 51; DB 6; Length 583;  
Best Local Similarity 100.0%; Pred. No. 6e-13;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2274 CTTCCAGCAATTTAGCATGATTCATGCAAGCTTGGAACTCTGGAGAGA 2324  
DB 439 CTTCCAGCAATTTAGCATGATTCATGCAAGCTTGGAACTCTGGAGAGA 389

RESULT 31  
ADR97461/c  
ID ADR97461 standard; DNA; 27 BP.

XX AC ADR97461;  
XX 02-DEC-2004 (first entry)  
XX Reverse quantitative PCR primer to amplify human CDIPT DNA.  
XX ss; apoptosis; cancer; inflammation; autoimmune;  
XX neurodegenerative disorder; cytostatic; antiinflammatory;  
XX immunosuppressive; neuroprotective; gene therapy; PCR; primer;  
XX real-time PCR; RT-PCR; CDIPT.

XX Homo sapiens.

XX WO2004078783-A2.

XX 16-SEP-2004.

XX 05-MAR-2004; 2004WO-GB000957.

XX 07-MAR-2003; 2003GB-00005267.

XX (EIRX-) EIRX THERAPEUTICS LTD.

XX Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;

XX WPI; 2004-662402/64.

XX Identifying an agent that modulates the function of an apoptosis-  
XX associated polypeptide, useful for diagnosing or treating e.g. cancer,  
XX comprises comparing the binding of the polypeptide to the candidate agent  
XX and to a control agent.

XX Example 5; Page 289; 304pp; English.

XX This invention relates to novel agents that modulates the function of  
XX human apoptosis-associated proteins specified within the specification.  
XX Specifically, it refers to a method for the identification of target  
XX genes whose expression is correlated with an early stage in the  
XX regulation of apoptosis. The present invention describes a method of  
XX contacting either candidate agents or control agents to the target genes  
XX and assessing the difference of binding and inhibitory activity, where  
XX the candidate agent is selected from a low molecular weight organic  
XX molecule, an antibody or fragment thereof, an antisense oligonucleotide,  
XX a small inhibitory dsRNA, or a ribozyme. As such, the compositions and  
XX methods are useful for diagnosing and treating diseases or conditions  
XX associated with abnormal apoptosis in mammalian tissue, such as cancer,  
XX inflammation, autoimmune or neurodegenerative disorders. Accordingly,  
XX they exhibit cytostatic, antiinflammatory, immunosuppressive and

CC neuroprotective activities. These may also be used for drug screening  
CC purposes and in gene therapy. This oligonucleotide sequence is a  
CC quantitative PCR primer used to amplify a target gene associated with  
CC apoptosis regulation, given in an exemplification of the invention.  
XX  
SQ Sequence 27 BP; 10 A; 11 C; 1 G; 5 T; 0 U; 0 Other;  
Query Match 1.2%; Score 27; DB 13; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 286 CTGATGGTGGAGATTTCGATGG 312  
DB 27 CTGATGGTGGAGATTTCGATGG 1  
RESULT 32  
ADR97432  
ID ADR97432 standard; DNA; 23 BP.  
XX  
AC ADR97432;  
DT 02-DEC-2004 (first entry)  
XX Forward quantitative PCR primer to amplify human Cathepsin C1 DNA.  
XX ss; apoptosis; cancer; inflammation; autoimmune;  
KW neurodegenerative disorder; cytostatic; antiinflammatory;  
KW immunosuppressive; neuroprotective; gene therapy; PCR; primer;  
KW real-time PCR; RT-PCR; Cathepsin C1.  
XX  
OS Homo sapiens.  
XX  
XX WO2004078783-A2.  
XX  
XX 16-SEP-2004.  
XX  
XX 05-MAR-2004; 2004WO-GB000957.  
XX  
XX 07-MAR-2003; 2003GB-00005267.  
XX  
XX (EIRX-) EIRX THERAPEUTICS LTD.  
XX  
XX Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;  
XX WPI; 2004-662402/64.  
XX  
XX Identifying an agent that modulates the function of an apoptosis-  
XX associated polypeptide, useful for diagnosing or treating e.g. cancer,  
XX comprises comparing the binding of the polypeptide to the candidate agent  
XX and to a control agent.  
XX  
XX Example 5; Page 289; 304pp; English.  
XX  
XX This invention relates to novel agents that modulates the function of  
XX human apoptosis-associated proteins specified within the specification.  
XX Specifically, it refers to a method for the identification of target  
XX genes whose expression is correlated with an early stage in the  
XX regulation of apoptosis. The present invention describes a method of  
XX contacting either candidate agents or control agents to the target genes  
XX and assessing the difference of binding and inhibitory activity, where  
XX the candidate agent is selected from a low molecular weight organic  
XX molecule, an antibody or fragment thereof, an antisense oligonucleotide,  
XX a small inhibitory dsRNA, or a ribozyme. As such, the compositions and  
XX methods are useful for diagnosing and treating diseases or conditions  
XX associated with abnormal apoptosis in mammalian tissue, such as cancer,  
XX inflammation, autoimmune or neurodegenerative disorders. Accordingly,  
XX they exhibit cytostatic, antiinflammatory, immunosuppressive and  
XX neuroprotective activities. These may also be used for drug screening  
XX purposes and in gene therapy. This oligonucleotide sequence is a  
XX quantitative PCR primer used to amplify a target gene associated with  
XX apoptosis regulation, given in an exemplification of the invention.  
XX

SQ Sequence 23 BP; 9 A; 5 C; 8 G; 1 T; 0 U; 0 Other;  
Query Match 1.0%; Score 23; DB 13; Length 23;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 106 CCGGAGAGACAAAGCATGGA 128  
DB 1 CCGGAGAGACAAAGCATGGA 23  
RESULT 33  
ABL25903  
ID ABL25903 standard; DNA; 855 BP.  
XX  
XX ABL25903;  
XX  
XX 26-MAR-2002 (first entry)  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 29182.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
XX OS  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
XX  
XX 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX  
XX Claim 1; SEQ ID NO 29182; 2lpp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
XX sequences (AB101840-AB16175) and the encoded proteins (ABBS7737-  
XX AB872072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 855 BP; 215 A; 293 C; 220 G; 127 T; 0 U; 0 Other;  
Query Match 0.9%; Score 22; DB 4; Length 855;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 36 AGCAGCGGCAACAGCAGCAGCA 57  
DB 139 AGCAGCGGCAACAGCAGCAGCA 160  
RESULT 34  
ABL19239  
ID ABL19239 standard; DNA; 1933 BP.  
XX

AC ABL19239;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9190.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PP 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
PS Claim 1; SEQ ID NO 9190; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1933 BP; 519 A; 591 C; 493 G; 330 T; 0 U; 0 Other;  
XX  
XX Query Match 0.9%; Score 22; DB 4; Length 1933;  
XX Best Local Similarity 100.0%; Pred. No. 18;  
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 40 GCGGCACAGCAGCAGCAGCGG 61  
XX  
DB 484 GCGGCACAGCAGCAGCAGCGG 505  
XX  
XX  
RESULT 35  
ABL09457  
ID ABL09457 standard; cDNA; 2227 BP.  
XX  
AC ABL09457;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 22853.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX

PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR P-PSDB; ABB65354.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
PS Claim 1; SEQ ID NO 22853; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2227 BP; 565 A; 720 C; 628 G; 314 T; 0 U; 0 Other;  
XX  
XX Query Match 0.9%; Score 22; DB 4; Length 2227;  
XX Best Local Similarity 100.0%; Pred. No. 18;  
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 38 CAGCGGCAACAGCAGCAGCAGC 59  
XX  
DB 1201 CAGCGGCAACAGCAGCAGCAGC 1222  
XX  
XX  
RESULT 36  
ABL25902/c  
ID ABL25902 standard; DNA; 2855 BP.  
XX  
AC ABL25902;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 29179.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PP 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.



```
XX PS Claim 1; SEQ ID NO 29179; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 2855 BP; 711 A; 668 C; 706 G; 770 T; 0 U; 0 Other;

Query Match 0.9%; Score 22; DB 4; Length 2855;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 AGCAGCGGCACAGCAGCAGCA 57
Db 1717 AGCAGCGGCACAGCAGCAGCA 1696

RESULT 37
ABL23387
ID ABL23387 standard; DNA; 3705 BP.
XX AC ABL23387;
XX XX
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21634.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 21634; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3705 BP; 940 A; 1012 C; 1047 G; 706 T; 0 U; 0 Other;

Query Match 0.9%; Score 22; DB 4; Length 2855;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 AGCAGCGGCACAGCAGCAGCA 57
Db 1717 AGCAGCGGCACAGCAGCAGCA 1696

RESULT 37
ABL23387
ID ABL23387 standard; DNA; 3705 BP.
XX AC ABL23387;
XX XX
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21634.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 21634; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3705 BP; 940 A; 1012 C; 1047 G; 706 T; 0 U; 0 Other;

Query Match 0.9%; Score 22; DB 4; Length 3705;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 AGCAGCGGCACAGCAGCAGCA 57
Db 276 AGCAGCGGCACAGCAGCAGCA 297

RESULT 38
ABL19238
ID ABL19238 standard; DNA; 3933 BP.
XX AC ABL19238;
XX XX
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9187.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 9187; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3933 BP; 1234 A; 995 C; 781 G; 923 T; 0 U; 0 Other;

Query Match 0.9%; Score 22; DB 4; Length 3933;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GCGGCAACAGCAGCAGCAGCG 61
Db 1484 GCGGCAACAGCAGCAGCAGCG 1505

RESULT 39
ABL09456
ID ABL09456 standard; cDNA; 3974 BP.
XX AC ABL09456;
```

```
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 22850.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI; 2001-656860/75.
XX DR P-PSDB; ABB65353.
XX DR 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 22850; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3974 BP; 1173 A; 1026 C; 915 G; 860 T; 0 U; 0 Other;

Query Match 0.9%; Score 22; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 CAGCGGCAACAGCAGCAGCAGC 59
Db 2201 CAGCGGCAACAGCAGCAGCAGC 2222

RESULT 40
ABL23386
ID ABL23386 standard; DNA; 5903 BP.
XX AC ABL23386;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21631.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PI Murphy FU, Sheehan DB, Keating KE, Hayes I, Seera L;
```

```
PF 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI; 2001-656860/75.
XX DR 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 21631; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 5903 BP; 1547 A; 1491 C; 1506 G; 1359 T; 0 U; 0 Other;

Query Match 0.9%; Score 22; DB 4; Length 5903;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 AGCAGCGGCAACAGCAGCAGCAGCA 57
Db 1276 AGCAGCGGCAACAGCAGCAGCA 1297

RESULT 41
ADR97567/c
ID ADR97567 standard; RNA; 21 BP.
XX AC ADR97567;
XX DT 02-DEC-2004 (first entry)
XX DE Antisense GMPR2 siRNA oligomer to knockout apoptosis related gene.
XX KW ss; DNA/RNA hybrid; apoptosis; cancer; inflammation; autoimmune;
XX KW neurodegenerative disorder; cytostatic; antiinflammatory;
XX KW immunosuppressive; neuroprotective; gene therapy; siRNA; GMPR2.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT misc_feature 20..21
XX FT /tag= a
XX FT /note= "Oligo dT bases added to the 3' end to stabilise
XX FT binding"
XX PN WO2004078783-A2.
XX PD 16-SEP-2004.
XX PF 05-MAR-2004; 2004WO-GB000957.
XX PR 07-MAR-2003; 2003GB-00005267.
XX PA (EIRX-) EIRX THERAPEUTICS LTD.
XX PI Murphy FU, Sheehan DB, Keating KE, Hayes I, Seera L;
```

XX WPI; 2004-662402/64.

XX Identifying an agent that modulates the function of an apoptosis-

XX associated polypeptide, useful for diagnosing or treating e.g. cancer,

XX comprises comparing the binding of the polypeptide to the candidate agent

XX and to a control agent.

XX Example 5; Page 293; 304pp; English.

XX This invention relates to novel agents that modulates the function of

XX human apoptosis-associated proteins specified within the specification.

XX Specifically, it refers to a method for the identification of target

XX genes whose expression is correlated with an early stage in the

XX regulation of apoptosis. The present invention describes a method of

XX contacting either candidate agents or control agents to the target genes

XX and assessing the difference of binding and inhibitory activity, where

XX the candidate agent is selected from a low molecular weight organic

XX molecule, an antibody or fragment thereof, an antisense oligonucleotide,

XX a small inhibitory dsRNA, or a ribozyme. As such, the compositions and

XX methods are useful for diagnosing and treating diseases or conditions

XX associated with abnormal apoptosis in mammalian tissue, such as cancer,

XX inflammation, autoimmune or neurodegenerative disorders. Accordingly,

XX they exhibit cytostatic, antiinflammatory, immunosuppressive and

XX neuroprotective activities. These may also be used for drug screening

XX purposes and in gene therapy. This oligonucleotide sequence is an siRNA

XX oligo used to knockout expression of a human gene identified to be

XX associated with the regulation of apoptosis, given in an exemplification

XX of the invention.

XX Sequence 21 BP; 4 A; 5 C; 3 G; 2 T; 7 U; 0 Other;

SQ

Query Match 0.9%; Score 21; DB 13; Length 21;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1787 AAACCTGGAGTTAATCCGAGA 1807

DB 21 AAACCTGGAGTTAATCCGAGA 1

RESULT 42

AAZ98400

ID AAZ98400 standard; DNA; 1845 BP.

XX

XX AAZ98400;

XX

XX 19-JUN-2000 (first entry)

XX Canine betal-adrenoceptor polynucleotide sequence.

XX

XX Betal-adrenoceptor; antisense; transcription; hypertension; canine;

XX myocardial ischemia; hypotensive; vasotropic; ds.

XX

XX Canis familiaris.

XX

XX WO200015783-A2.

XX

XX 23-MAR-2000.

XX

XX 14-SEP-1999; 99WO-US021007.

XX

XX 14-SEP-1998; 98US-00152717.

XX

XX (UYFL ) UNIV FLORIDA.

XX

XX Phillips MI, Zhang Y;

XX

XX WPI; 2000-271424/23.

XX

XX New antisense oligonucleotides, useful for treatment and prevention of

XX hypertension and myocardial ischemia, are directed against mRNA encoding

XX the betal-adrenoceptor.

PT

XX

PS

XX Example 4; Page 90-91; 144pp; English.

XX The invention provides a composition comprising at least one

XX oligonucleotide, of 9-35 bases, that binds specifically to part of the

XX mRNA, expressed from the gene encoding a mammalian betal-adrenoceptor

XX polypeptide. Binding of the oligo to the betal-adrenoceptor mRNA alters

XX transcription of the mRNA in cells. The compositions are used to treat

XX diseases associated with increased number or activity of the betal-

XX adrenoceptor polypeptide, specifically hypertension and myocardial

XX ischemia. The compositions can also be used to produce recombinant

XX (viral) vectors for use in vaccines, for treating or preventing the

XX diseases. The antisense oligos may also be used to detect the mRNA or

XX related DNA, including visualization within a cell, to generate

XX transgenic animals with altered betal-adrenoceptor activity, and to

XX screen patients for susceptibility to hypertension, by detection of

XX particular alleles of betal-adrenoceptor genes. The present sequence

XX represents the polynucleotide sequence of the canine betal-adrenoceptor

XX (GenBank Accn No: U73207)

XX Sequence 1845 BP; 197 A; 754 C; 647 G; 246 T; 0 U; 1 Other;

SQ

Query Match 0.9%; Score 21; DB 3; Length 1845;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CAGCAGCGGCTCCGCGAGCGG 73

DB 93 CAGCAGCGGCTCCGCGAGCGG 113

RESULT 43

ABK40732

ID ABK40732 standard; DNA; 1845 BP.

XX

XX ABK40732;

XX

XX 21-MAY-2002 (first entry)

XX Dog betal-adrenoceptor gene.

XX

XX ds; antisense; betal adrenoceptor; betal-AR; vasotropic; hypotensive;

XX cardiant; hypertension; hypertrophy; cardiac ischaemia;

XX cardiovascular disease; cardiac dysfunction; gene.

XX

XX Canis familiaris.

XX

XX WO200204623-A2.

XX

XX 17-JAN-2002.

XX

XX 11-JUL-2001; 2001WO-US021759.

XX

XX 11-JUL-2000; 2000US-00614034.

XX

XX (UYFL ) UNIV FLORIDA.

XX

XX Phillips MI, Zhang Y;

XX

XX WPI; 2002-164644/21.

XX

XX Novel antisense oligonucleotides that specifically bind to mRNA encoding

XX beta l-adrenoceptor polypeptide, useful for treating cardiac

XX dysfunction, hypertension, hypertrophy and other cardiovascular diseases

XX in humans.

XX

XX Example 8; Page 84-85; 186pp; English.

XX

XX The invention relates to an isolated antisense oligonucleotide of 9-35

XX nucleotides in length, which specifically binds to a portion of an mRNA

XX expressed from a gene encoding a mammalian betal-adrenoceptor (AR)

XX polypeptide and alters the translation of the mRNA into the betal-AR

XX polypeptide in a host cell expressing the mRNA. Also included are a

XX

CC recombinant vector comprising the antisense oligonucleotide, and a host  
CC cell comprising the vector. A composition comprising the antisense  
CC oligonucleotides is useful in the manufacture of a medicament for use in  
CC treating or ameliorating hypertension, hypertrophy and cardiac ischaemia  
CC in a mammal. A composition comprising the antisense oligonucleotides is  
CC also useful for reducing the level of beta-AR polypeptide, the  
CC transcription of beta-AR polypeptide-specific mRNA in a mammalian host  
CC cell, particularly human cell, and for decreasing blood pressure in a  
CC mammal, where the antisense oligonucleotide is operably linked to a  
CC promoter capable of expressing the oligonucleotide in the cell. A  
CC composition comprising a selected nucleic acid segment that comprises a  
CC full-length, or is a full length betaal-adrenoceptor antisense  
CC polynucleotide operatively linked to a promoter capable of expressing the  
CC polynucleotide in a cell is also useful for reducing the level of beta-1-  
CC adrenoceptor polypeptide in a mammalian host cell. The antisense  
CC oligonucleotide is also useful for other cardiovascular diseases and  
CC cardiac dysfunction in humans. The present sequence is a mammalian beta-  
CC AR gene, the mRNA for which is a target for antisense oligonucleotides of  
CC the invention

XX SQ Sequence 1845 BP; 197 A; 754 C; 647 G; 246 T; 0 U; 1 Other;

Query Match 0.9%; Score 21; DB 6; Length 1845;  
Best Local Similarity 100.0%; Pred. NO. 53;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CAGCAGCGGCTCCGCGAGCG 73  
DB 93 CAGCAGCGGCTCCGCGAGCG 113

RESULT 44

AAAX60124  
ID AAX60124 standard; CDNA; 4011 BP.

XX AC AAX60124;

XX DT 05-AUG-1999 (first entry)

XX DE DNA sequence of the 5' region/exon 1/intron 1 of murine SAP gene.

XX KW Signalling lymphocyte activation molecule; SLAM; SLAM associated protein;  
XX SAP; EAT-2; antigen-specific T cell activation;  
XX X-linked proliferative disease; XLP; tuberculosis; AIDS; schistosomiasis;  
XX cancer; breast cancer; prostate cancer; leukemia; leprosy;  
XX autoimmune disease; ss.

XX OS Mus musculus.

XX PN W09926980-A1.

XX PD 03-JUN-1999.

XX PF 19-NOV-1998; 98WO-US024976.

XX PR 21-NOV-1997; 97US-00376096.

XX PR 04-SEP-1998; 98US-0039160P.

XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX PI Terhorst CP, Sayos-Ortega J;

XX PS WPI; 1999-357816/30.

XX PT New signaling lymphocyte activation molecule associated protein useful in  
XX the treatment of tuberculosis and AIDS.

XX PS Disclosure; Fig 15; 124pp; English.

XX CC The present sequence represents part of a signalling lymphocyte  
XX activation molecule (SLAM) associated protein (SAP) gene. A SAP family  
XX member polypeptide (such as EAT-2) or a fragment, mutant or fusion can be  
XX used for treating a disease involving aberrant SH2 domain-containing

CC protein-mediated signal transduction in a patient. Upregulation of SAP  
CC family members can be used for the treatment of diseases involving  
CC insufficient antigen-specific T cell activation (e.g. X-linked  
CC proliferative disease (XLP), tuberculosis, AIDS, schistosomiasis) or  
CC cancer (e.g. breast cancer, prostate cancer, leukemia), or leprosy.  
CC Downregulation of SAP family members can be used for the treatment of  
CC diseases characterized by an excessive amount of antigen-specific T cell  
CC activation (e.g. autoimmune diseases). The products can also be used for  
CC detection and diagnosis

XX SQ Sequence 4011 BP; 1125 A; 828 C; 885 G; 1173 T; 0 U; 0 Other;

Query Match 0.9%; Score 21; DB 2; Length 4011;  
Best Local Similarity 100.0%; Pred. NO. 51;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GCAGCGGCGACAGCAGCAGCA 57  
DB 3157 GCAGCGGCGACAGCAGCAGCA 3177

RESULT 45

AAF89814/c

ID AAF89814 standard; DNA; 20 BP.

XX AC AAF89814;

XX DT 23-JUL-2001 (first entry)

XX DE Human aspartyl (asparaginyl) beta-hydroxylase antisense molecule.

XX KW Epidermal growth factor-like domain; EGF-like domain; cancer;  
XX human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour;  
XX antisense oligonucleotide; ss.

XX OS Homo sapiens.

XX PN W0200135102-A2.

XX PD 17-MAY-2001.

XX PF 08-NOV-2000; 2000WO-US030738.

XX PR 08-NOV-1999; 99US-00436184.

XX PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.

XX PI Wands JR, De La Monte SM, Ince N, Carlson RI;

XX PS WPI; 2001-329171/34.

XX PT Diagnosing malignant neoplasm in a mammal, involves contacting mammalian  
XX sample with antibody that binds to human aspartyl beta-hydroxylase  
XX polypeptide to form antigen-antibody complex and detecting the complex.

XX PS Claim 16; Page 57; 76pp; English.

XX CC The present sequence represents an antisense oligonucleotide which  
XX inhibits human aspartyl (asparaginyl) beta-hydroxylase (HAAH) gene.  
XX Epidermal growth factor (EGF)-like domains of polypeptides are  
XX hydroxylated by HAAH enzymes. HAAH is used in the method of the  
XX invention. The specification describes a method for diagnosing a  
XX malignant neoplasm in a mammal. The method comprises contacting a body  
XX fluid with an antibody which binds to HAAH polypeptide under complex  
XX forming conditions, and detecting the antigen-antibody complex. The  
XX method is useful for diagnosing and prognosing a malignant neoplasm in a  
XX bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
XX serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
XX where the neoplasm is derived from endodermal tissue and is selected from  
XX colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
XX bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
XX agent, are useful for killing tumour cells

SQ Sequence 20 BP; 2 A; 7 C; 7 G; 4 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATGGCCAG 20  
DB 20 CGGACCGTGCATGGCCAG 1  
RESULT 46  
AAF89813/c  
ID AAF89813 standard; DNA; 20 BP.  
XX  
AC AAF89813;  
XX  
DT 23-JUL-2001 (first entry)  
XX  
DE Human aspartyl (asparaginyl) beta-hydroxylase antisense molecule.  
XX  
KW Epidermal growth factor-like domain; EGF-like domain; cancer;  
KW human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour;  
KW antisense oligonucleotide; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200135102-A2.  
XX  
PD 17-MAY-2001.  
XX  
PF 08-NOV-2000; 2000WO-US030738.  
XX  
PR 08-NOV-1999; 99US-00436184.  
XX  
PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.  
XX  
PI Wands JR, De La Monte SM, Ince N, Carlson RI;  
XX WPI; 2001-329171/34.  
XX  
PT Diagnosing malignant neoplasm in a mammal, involves contacting mammalian  
PT sample with antibody that binds to human aspartyl beta-hydroxylase  
PT polypeptide to form antigen-antibody complex and detecting the complex.  
XX  
PS Claim 14; Page 57; 76pp; English.  
XX  
CC The present sequence represents an antisense oligonucleotide which  
CC inhibits human aspartyl (asparaginyl) beta-hydroxylase (HAAH) gene.  
CC Epidermal growth factor (EGF)-like domains of polypeptides are  
CC hydroxylated by HAAH enzymes. HAAH is used in the method of the  
CC invention. The specification describes a method for diagnosing a  
CC malignant neoplasm in a mammal. The method comprises contacting a body  
CC fluid with an antibody which binds to HAAH polypeptide under complex  
CC forming conditions, and detecting the antigen-antibody complex. The  
CC method is useful for diagnosing and prognosing a malignant neoplasm in a  
CC bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
CC serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
CC where the neoplasm is derived from endodermal tissue and is selected from  
CC colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
CC bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
CC agent, are useful for killing tumour cells  
SQ Sequence 20 BP; 3 A; 6 C; 4 G; 7 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 AATGCCCGCAGCGTAAGATG 30  
DB 20 AATGCCCGCAGCGTAAGATG 1  
RESULT 48  
ABS57871/c  
ID ABS57871 standard; DNA; 20 BP.  
XX  
AC ABS57871;  
XX  
DT 06-FEB-2003 (first entry)  
XX  
DE Human HAAH antisense oligonucleotide #3.

SQ Sequence 20 BP; 2 A; 7 C; 7 G; 4 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATGGCCAG 20  
DB 20 CGGACCGTGCATGGCCAG 1  
RESULT 46  
AAF89813/c  
ID AAF89813 standard; DNA; 20 BP.  
XX  
AC AAF89813;  
XX  
DT 23-JUL-2001 (first entry)  
XX  
DE Human aspartyl (asparaginyl) beta-hydroxylase antisense molecule.  
XX  
KW Epidermal growth factor-like domain; EGF-like domain; cancer;  
KW human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour;  
KW antisense oligonucleotide; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200135102-A2.  
XX  
PD 17-MAY-2001.  
XX  
PF 08-NOV-2000; 2000WO-US030738.  
XX  
PR 08-NOV-1999; 99US-00436184.  
XX  
PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.  
XX  
PI Wands JR, De La Monte SM, Ince N, Carlson RI;  
XX WPI; 2001-329171/34.  
XX  
PT Diagnosing malignant neoplasm in a mammal, involves contacting mammalian  
PT sample with antibody that binds to human aspartyl beta-hydroxylase  
PT polypeptide to form antigen-antibody complex and detecting the complex.  
XX  
PS Claim 15; Page 57; 76pp; English.  
XX  
CC The present sequence represents an antisense oligonucleotide which  
CC inhibits human aspartyl (asparaginyl) beta-hydroxylase (HAAH) gene.  
CC Epidermal growth factor (EGF)-like domains of polypeptides are  
CC hydroxylated by HAAH enzymes. HAAH is used in the method of the  
CC invention. The specification describes a method for diagnosing a  
CC malignant neoplasm in a mammal. The method comprises contacting a body  
CC fluid with an antibody which binds to HAAH polypeptide under complex  
CC forming conditions, and detecting the antigen-antibody complex. The  
CC method is useful for diagnosing and prognosing a malignant neoplasm in a  
CC bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
CC serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
CC where the neoplasm is derived from endodermal tissue and is selected from  
CC colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
CC bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
CC agent, are useful for killing tumour cells  
SQ Sequence 20 BP; 3 A; 6 C; 4 G; 5 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 CGTGCATGGCCAGCGTAA 25  
DB 20 CGTGCATGGCCAGCGTAA 1

XX Human; ss; aspartyl (asparaginyl) beta-hydroxylase; HAAH; antisense;  
 KW cytosolic; immunostimulant; antibody; neoplasm; tumour; FB50; 86A; 5C7;  
 KW 19B; brain tumour; glioma; glioblastoma; astrocytoma; haemangioma;  
 KW pancreatic carcinoma; colon cancer; breast cancer; pancreatic cancer;  
 KW liver cancer; cancer of the bile ducts; primary malignant CNS neoplasm;  
 KW metastatic CNS neoplasm.  
 XX Homo sapiens.  
 XX US2002110559-A1.  
 XX 15-AUG-2002.  
 XX 17-MAY-2001; 2001US-00859604.  
 XX 08-NOV-1999; 99US-00436184.  
 XX (WAND/) WANDS J R.  
 XX (DMON/) DE LA MONTE S M.  
 XX (DEUT/) DEUTCH A H.  
 XX (GHAN/) GHANBARI H A.  
 XX Wands JR, De La Monte SM, Deutch AH, Ghanbari HA;  
 XX WPI; 2003-066676/06.  
 XX Diagnosing neoplasm and inhibiting tumor growth, by contacting tissue of  
 PT mammal with detectably-labeled antibody which binds to human aspartyl  
 PT (asparaginyl) beta-hydroxylase.  
 XX Example 6; Page 18; 34pp; English.  
 XX The invention relates to diagnosing a neoplasm and inhibiting tumor  
 CC growth in a mammal, using an antibody that binds to human aspartyl  
 CC (asparaginyl) beta-hydroxylase (HAAH). Diagnosing a neoplasm comprises  
 CC contacting a tissue with a detectably-labelled antibody where an increase  
 CC in level of antibody binding at tissue site compared to the level of  
 CC binding to normal non-neoplastic tissue indicates the presence of a  
 CC neoplasm at the tissue site. Inhibiting tumor growth in mammal involves  
 CC administering the antibody conjugated to a cytotoxic agent to a mammal.  
 CC Also included are a method of conferring an immune response to a tumour  
 CC cell in a mammal, by administering the antibody, a method of inducing an  
 CC HAAH-specific immune response in a mammal, by administering to the mammal  
 CC an HAAH polypeptide (or a polynucleotide composition encoding the  
 CC polypeptide, or its degenerate variant), a fragment of HAAH comprising an  
 CC extracellular domain and lacking a cytoplasmic domain of HAAH, an  
 CC antibody or its fragment which binds to HAAH (where the antibody is FB50,  
 CC 86A, 5C7 or 19B), a hybridoma cell line chosen from hybridoma FB50,  
 CC HA386A, HA15CA and HA219B, and a fragment of HAAH which lacks enzymatic  
 CC activity or alpha-ketoglutarate binding domain and epidermal growth  
 CC factor (EGF)-like domain. The methods are useful for diagnosing neoplasm  
 CC in a mammal, inhibiting tumour growth in a mammal, conferring an immune  
 CC response to a brain tumour cell (e.g. glioma, glioblastoma, astrocytoma  
 CC or haemangioma) in a mammal, for conferring an immune response to a  
 CC pancreatic carcinoma cell and for inducing a HAAH-specific immune  
 CC response in a mammal. The method is useful for diagnosing malignant  
 CC neoplasms derived from endodermal tissue, e.g. colon cancer, breast  
 CC cancer, pancreatic cancer, liver cancer and cancer of the bile ducts. The  
 CC method is also useful for diagnosing neoplasms of central nervous system  
 CC (CNS) e.g. primary malignant CNS neoplasms of both neuronal and glial  
 CC cell origin and metastatic CNS neoplasms, and for diagnosing brain  
 CC tumours e.g. glioma, glioblastoma, astrocytoma or haemangioma. The  
 CC present sequence represents a Human HAAH antisense oligonucleotide  
 XX Sequence 20 BP; 2 A; 7 C; 7 G; 4 T; 0 U; 0 Other;  
 SQ  
 Query Match 0.9%; Score 20; DB 8; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 CGGACCGTGCATGGCCAG 20  
 |||||

Db 20 CGGACCGTGCATGGCCAG 1  
 RESULT 49  
 ABS57870/c  
 ID ABS57870 standard; DNA; 20 BP.  
 XX ABS57870;  
 XX 06-FEB-2003 (first entry)  
 XX Human HAAH antisense oligonucleotide #2.  
 XX Human; ss; aspartyl (asparaginyl) beta-hydroxylase; HAAH; antisense;  
 KW cytosolic; immunostimulant; antibody; neoplasm; tumour; FB50; 86A; 5C7;  
 KW 19B; brain tumour; glioma; glioblastoma; astrocytoma; haemangioma;  
 KW pancreatic carcinoma; colon cancer; breast cancer; pancreatic cancer;  
 KW liver cancer; cancer of the bile ducts; primary malignant CNS neoplasm;  
 KW metastatic CNS neoplasm.  
 XX Homo sapiens.  
 XX US2002110559-A1.  
 XX 15-AUG-2002.  
 XX 17-MAY-2001; 2001US-00859604.  
 XX 08-NOV-1999; 99US-00436184.  
 XX (WAND/) WANDS J R.  
 XX (DMON/) DE LA MONTE S M.  
 XX (DEUT/) DEUTCH A H.  
 XX (GHAN/) GHANBARI H A.  
 XX Wands JR, De La Monte SM, Deutch AH, Ghanbari HA;  
 XX WPI; 2003-066676/06.  
 XX Diagnosing neoplasm and inhibiting tumor growth, by contacting tissue of  
 PT mammal with detectably-labeled antibody which binds to human aspartyl  
 PT (asparaginyl) beta-hydroxylase.  
 XX Example 6; Page 18; 34pp; English.  
 XX The invention relates to diagnosing a neoplasm and inhibiting tumor  
 CC growth in a mammal, using an antibody that binds to human aspartyl  
 CC (asparaginyl) beta-hydroxylase (HAAH). Diagnosing a neoplasm comprises  
 CC contacting a tissue with a detectably-labelled antibody where an increase  
 CC in level of antibody binding at tissue site compared to the level of  
 CC binding to normal non-neoplastic tissue indicates the presence of a  
 CC neoplasm at the tissue site. Inhibiting tumor growth in mammal involves  
 CC administering the antibody conjugated to a cytotoxic agent to a mammal.  
 CC Also included are a method of conferring an immune response to a tumour  
 CC cell in a mammal, by administering the antibody, a method of inducing an  
 CC HAAH-specific immune response in a mammal, by administering to the mammal  
 CC an HAAH polypeptide (or a polynucleotide composition encoding the  
 CC polypeptide, or its degenerate variant), a fragment of HAAH comprising an  
 CC extracellular domain and lacking a cytoplasmic domain of HAAH, an  
 CC antibody or its fragment which binds to HAAH (where the antibody is FB50,  
 CC 86A, 5C7 or 19B), a hybridoma cell line chosen from hybridoma FB50,  
 CC HA386A, HA15CA and HA219B, and a fragment of HAAH which lacks enzymatic  
 CC activity or alpha-ketoglutarate binding domain and epidermal growth  
 CC factor (EGF)-like domain. The methods are useful for diagnosing neoplasm  
 CC in a mammal, inhibiting tumour growth in a mammal, conferring an immune  
 CC response to a brain tumour cell (e.g. glioma, glioblastoma, astrocytoma  
 CC or haemangioma) in a mammal, for conferring an immune response to a  
 CC pancreatic carcinoma cell and for inducing a HAAH-specific immune  
 CC response in a mammal. The method is useful for diagnosing malignant  
 CC neoplasms derived from endodermal tissue, e.g. colon cancer, breast  
 CC cancer, pancreatic cancer, liver cancer and cancer of the bile ducts. The  
 CC method is also useful for diagnosing neoplasms of central nervous system  
 CC (CNS) e.g. primary malignant CNS neoplasms of both neuronal and glial  
 CC cell origin and metastatic CNS neoplasms, and for diagnosing brain  
 CC tumours e.g. glioma, glioblastoma, astrocytoma or haemangioma. The  
 CC present sequence represents a Human HAAH antisense oligonucleotide  
 XX Sequence 20 BP; 2 A; 7 C; 7 G; 4 T; 0 U; 0 Other;  
 SQ

CC cell origin and metastatic CNS neoplasms, and for diagnosing brain  
CC tumours e.g. glioma, glioblastoma, astrocytoma or haemangioma. The  
CC present sequence represents a Human HAAH antisense oligonucleotide  
XX  
SQ Sequence 20 BP; 3 A; 6 C; 6 G; 5 T; 0 U; 0 Other;  
  
Query Match 0.9%; Score 20; DB 8; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 6 CGTCAATGGCCAGCGTAA 25  
Db 20 CGTCAATGGCCAGCGTAA 1  
  
RESULT 50  
ABS57869/c  
ID ABS57869 standard; DNA; 20 BP.  
XX  
AC ABS57869;  
XX  
DT 06-FEB-2003 (first entry)  
XX  
DE Human HAAH antisense oligonucleotide #1.  
XX  
KW Human; ss; aspartyl (asparaginyl) beta-hydroxylase; HAAH; antisense;  
KW cytosolic; immunostimulant; antibody; neoplasm; tumour; FB50; 86A; 5C7;  
KW 19B; brain tumour; glioma; glioblastoma; astrocytoma; haemangioma;  
KW pancreatic carcinoma; colon cancer; breast cancer; pancreatic cancer;  
KW liver cancer; cancer of the bile ducts; primary malignant CNS neoplasm;  
KW metastatic CNS neoplasm.  
XX  
KW Homo sapiens.  
XX  
OS  
XX  
FN US2002110559-A1.  
XX  
PD 15-AUG-2002.  
XX  
PF 17-MAY-2001; 2001US-00859604.  
XX  
PR 08-NOV-1999; 99US-00436184.  
XX  
PA (WAND/) WANDS J R.  
PA (DMON/) DE LA MONTE S M.  
PA (DEUT/) DEUTCH A H.  
PA (GHAN/) GHANBARI H A.  
XX  
PI Wands JR, De La Monte SM, Deutch AH, Ghanbari HA;  
XX  
DR WPI; 2003-066676/06.  
XX  
PT Diagnosing neoplasm and inhibiting tumor growth, by contacting tissue of  
PT mammal with detectably-labeled antibody which binds to human aspartyl  
PT (asparaginyl) beta-hydroxylase.  
XX  
PS Example 6; Page 18; 34pp; English.  
XX  
CC The invention relates to diagnosing a neoplasm and inhibiting tumour  
CC growth in a mammal, using an antibody that binds to human aspartyl  
CC (asparaginyl) beta-hydroxylase (HAAH). Diagnosing a neoplasm comprises  
CC contacting a tissue with a detectably-labelled antibody where an increase  
CC in level of antibody binding at tissue site compared to the level of  
CC binding to normal non-neoplastic tissue indicates the presence of a  
CC neoplasm at the tissue site. Inhibiting tumour growth in mammal involves  
CC administering the antibody conjugated to a cytotoxic agent to a mammal.  
CC Also included are a method of conferring an immune response to a tumour  
CC cell in a mammal, by administering the antibody, a method of inducing an  
CC HAAH-specific immune response in a mammal, by administering to the mammal  
CC an HAAH polypeptide (or a polynucleotide composition encoding the  
CC polypeptide, or its degenerate variant), a fragment of HAAH comprising an  
CC extracellular domain and lacking a cytoplasmic domain of HAAH, an  
CC antibody or its fragment which binds to HAAH (where the antibody is FB50,  
CC 86A, 5C7 or 19B), a hybridoma cell line chosen from hybridoma FB50,

CC HA386A, HA15CA and HA219B, and a fragment of HAAH which lacks enzymatic  
CC activity or alpha-ketoglutarate binding domain and epidermal growth  
CC factor (EGF)-like domain. The methods are useful for diagnosing neoplasm  
CC in a mammal, inhibiting tumour growth in a mammal, conferring an immune  
CC response to a brain tumour cell (e.g. glioma, glioblastoma, astrocytoma  
CC or haemangioma) in a mammal, for conferring immune response to a  
CC pancreatic carcinoma cell and for inducing a HAAH-specific immune  
CC response in a mammal. The method is useful for diagnosing malignant  
CC neoplasms derived from endodermal tissue, e.g. colon cancer, breast  
CC cancer, pancreatic cancer, liver cancer and cancer of the bile ducts. The  
CC method is also useful for diagnosing neoplasms of both nervous system  
CC (CNS) e.g. primary malignant CNS neoplasms, and for diagnosing brain  
CC cell origin and metastatic CNS neoplasms, and for diagnosing brain  
CC tumours e.g. glioma, glioblastoma, astrocytoma or haemangioma. The  
CC present sequence represents a Human HAAH antisense oligonucleotide  
XX  
SQ Sequence 20 BP; 3 A; 6 C; 4 G; 7 T; 0 U; 0 Other;  
  
Query Match 0.9%; Score 20; DB 8; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 11 AATGGCCCGCGTAAGAATG 30  
Db 20 AATGGCCCGCGTAAGAATG 1  
  
RESULT 51  
ADR37507  
ID ADR37507 standard; RNA; 21 BP.  
XX  
AC ADR97507;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Antisense PPIA siRNA oligomer to knockout apoptosis related gene.  
XX  
KW ss; DNA/RNA hybrid; apoptosis; cancer; inflammation; autoimmune;  
KW neurodegenerative disorder; cytostatic; antiinflammatory; PPIA.  
KW immunosuppressive; neuroprotective; gene therapy; siRNA; PPIA.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 20..21 /\*tag= a  
FT binding /note= "Oligo dT bases added to the 3' end to stabilise  
XX binding"  
XX  
XX WO2004078783-A2.  
XX  
XX 16-SEP-2004.  
XX  
XX 05-MAR-2004; 2004WO-GB000957.  
XX  
XX 07-MAR-2003; 2003GB-00005267.  
XX  
XX (EIRX-) EIRX THERAPEUTICS LTD.  
XX  
XX Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;  
XX WPI; 2004-662402/64.  
XX  
XX Identifying an agent that modulates the function of an apoptosis-  
XX associated polypeptide, useful for diagnosing or treating e.g. cancer,  
XX comprises comparing the binding of the polypeptide to the candidate agent  
XX and to a control agent.  
XX  
XX Example 5; Page 293; 304pp; English.  
XX  
XX This invention relates to novel agents that modulates the function of  
XX human apoptosis-associated proteins specified within the specification.  
CC

Specifically, it refers to a method for the identification of target genes whose expression is correlated with an early stage in the regulation of apoptosis. The present invention describes a method of contacting either candidate agents or control agents to the target genes and assessing the difference of binding and inhibitory activity, where the candidate agent is selected from a low molecular weight organic molecule, an antibody or fragment thereof, an antisense oligonucleotide, a small inhibitory dsRNA, or a ribozyme. As such, the compositions and methods are useful for diagnosing and treating diseases or conditions associated with abnormal apoptosis in mammalian tissues, such as cancer, inflammation, autoimmune or neurodegenerative disorders. Accordingly, they exhibit cytostatic, antiinflammatory, immunosuppressive and neuroprotective activities. These may also be used for drug screening purposes and in gene therapy. This oligonucleotide sequence is an siRNA oligo used to knockout expression of a human gene identified to be associated with the regulation of apoptosis, given in an exemplification of the invention.

Sequence 21 BP; 7 A; 3 C; 5 G; 2 T; 4 U; 0 Other;

Query Match 0.9%; Score 20; DB 13; Length 21;  
Best Local Similarity 80.0%; Pred. No. 1-7e+02;  
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1789 ACTGGAGTATCCGAGAT 1808  
DB 1 ACUGGAGUUAUCCGAGAT 20

RESULT 52  
ACH75949/c  
ID ACH75949 standard; DNA; 550 BP.  
XX ACH75949;  
XX 29-JUL-2004 (first entry)  
XX Human genome derived single exon probe #9144.  
XX Human; probe; ss; gene expression; single exon probe; microarray;  
XX alternative splicing event; genomic alteration.

OS Homo sapiens.  
XX US2003194704-A1.  
XX 16-OCT-2003.  
XX 03-APR-2002; 2002US-00029386.  
XX 03-APR-2002; 2002US-00029386.  
XX (PENN/) PENN S G.  
XX (RANK/) RANK D R.  
XX (HANZ/) HANZEL D K.  
XX Penn SG, Rank DR, Hanzel DK;  
XX WPI; 2004-119264/12.  
XX New human genome-derived single exon nucleic acid probes useful for human gene expression analysis, for identifying or characterizing alternative splicing events, for assessing genomic alterations or as tools for surveying tissues.  
XX Claim 15; SEQ ID NO 9144; 80pp; English.

The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 688 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule

expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality, a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704

Sequence 550 BP; 144 A; 121 C; 147 G; 138 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 12; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1256 TGCAGACCTGCTGAAGCTGA 1275  
DB 280 TGCAGACCTGCTGAAGCTGA 261

RESULT 53  
ACH96138/c  
ID ACH96138 standard; DNA; 786 BP.  
XX ACH96138;  
XX 29-JUL-2004 (first entry)  
XX Klebsiella pneumoniae polynucleotide seqid 1933.  
XX Recombinant expression vector; transcription regulatory element;  
XX Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.  
XX Klebsiella pneumoniae.  
XX US6610836-B1.  
XX 26-AUG-2003.  
XX 27-JAN-2000; 2000US-00489039.  
XX 29-JAN-1999; 99US-0117747P.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Breton GL, Osborne M;  
XX WPI; 2003-895346/82.  
XX P-PSDB; ABO62587.  
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for preparing a vaccine composition against Klebsiella pneumoniae.



```
XX PS Disclosure; SEQ ID NO 1933; 932pp; English.
XX CC
XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
XX CC pneumoniae polypeptide. Also described are: a recombinant expression
XX CC vector comprising the nucleic acid, operably linked to a transcription
XX CC regulatory element; and a cell comprising the recombinant expression
XX CC vector. The nucleic acid is useful for preparing a vaccine composition
XX CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
XX CC pneumoniae polypeptide of the invention
XX SQ Sequence 786 BP; 129 A; 220 C; 235 G; 202 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 11; Length 786;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 38 CAGCGGCAACAGCAGCAGCA 57
DB 294 CAGCGGCAACAGCAGCAGCA 275
RESULT 54
AAT62523
ID AAT62523 standard; DNA; 849 BP.
XX AC
XX AAT62523;
XX DT 06-JUN-1997 (first entry)
XX DE IGSP-hPOMC fusion DNA sequence.
XX KW Analgesic; pain; bioartificial organ; pro-opiomelanocotin; POMC;
XX KW beta-endorphin; CNTF; IGSP-hPOMC; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT exon 1..89 /tag= b
FT 5'UTR 1..32 /tag= a
FT misc_feature 43..186 /tag= f
FT /product= "IGSP region"
FT intron 90..168
FT exon 169..849 /tag= c
FT misc_feature 187..806 /tag= d
FT 3'UTR 807..849 /tag= e
FT /product= "hPOMC region"
XX PN WO9640959-A1.
XX XX
XX XX 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009629.
XX XX
XX PF 07-JUN-1995; 95US-00481917.
XX XX
XX PA (CYTO-) CYTOTHERAPEUTICS INC.
XX XX Saydoff J, Wong S;
XX XX WPI; 1997-087062/08.
XX XX
XX XX Stably transformed cells expressing endorphin, enkephalin and
XX PT catecholamine - and artificial organs contg. them, useful for control of
XX PT pain, esp. implanted in the CNS.
XX XX
```

```
PS Example; Page 58; 114pp; English.
XX CC
XX CC A DNA sequence (AAT62523) comprising a fusion between IGSP and human pro-
XX CC opiomelanocortin (POMC) DNA can be utilised in the construction of
XX CC vectors for the expression of endorphin cpds. Cleavage enzymes present in
XX CC host cells can generate a suite of endorphins from POMC, some or all of
XX CC which have analgesic properties. A DNA sequence (see also AAT62524) in
XX CC which the ACTH coding region of the POMC gene is deleted is preferred.
XX CC Host cells transfected with vectors (or a single polycistronic vector)
XX CC encoding analgesic cpds. can be encapsulated to form bioartificial organs
XX CC that are implanted into a patient for control of pain
XX SQ Sequence 849 BP; 176 A; 266 C; 286 G; 121 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 2; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GCAACAGCAGCAGCAGCGGC 62
DB 271 GCAACAGCAGCAGCAGCGGC 290
RESULT 55
ABA19178/c
ID ABA19178 standard; DNA; 900 BP.
XX AC
XX ABA19178;
XX DT 23-JAN-2002 (first entry)
XX DE Human nervous system related polynucleotide SEQ ID NO 11509.
XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
XX KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
XX KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX OS Homo sapiens.
XX PN WO200159063-A2.
XX PD 16-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001334.
XX XX
XX XX 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-019076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-021680P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
```



Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0; Mismatches 0; Conservative 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62  
Db 764 GCAACAGCAGCAGCAGCGGC 745

RESULT 56  
AAC58389  
ID AAC58389 standard; cDNA; 908 BP.

XX AC AAC58389;

XX DT 29-JAN-2001 (first entry)

XX DE Human PRO23 nucleotide sequence SEQ ID NO: 64.

XX KW Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth; proliferation; tumorigenesis; identification; cancer; cytostatic; neurotropic; neuroprotective; antiinflammatory; immunosuppressive; immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy; neuronal disorder; glioma disorder; astrocytoma disorder; angiogenic; hypothalamic disorder; glandular disorder; macrophagal disorder; epithelial disorder; stromal disorder; blastocoele disorder; inflammatory disorder; immunologic disorder; ss.

XX OS Homo sapiens.

XX PN WO200053755-A2.

XX PD 14-SEP-2000.

XX PF 06-JAN-2000; 2000WO-US000376.

XX PR 08-MAR-1999; 99WO-US005028.

XX PR 02-JUN-1999; 99WO-US012252.

XX PR 23-JUN-1999; 98US-0141037P.

XX PR 07-JUL-1999; 98US-0143048P.

XX PR 26-JUL-1999; 99US-0145698P.

XX PR 30-NOV-1999; 99WO-US028313.

XX PR 20-DEC-1999; 99WO-US030911.

XX PR 05-JAN-2000; 2000WO-US000219.

XX PA (GETH ) GENENTECH INC.

XX PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillian KJ, Roy MA; Watanabe CK, Wood WI;

XX P-PSDB; AAB24079.

XX WPI; 2000-572270/53.

XX DR Thirty PRO polynucleotides encoding PRO polypeptides, useful in the treatment, diagnosis and prevention of cancer.

XX PT Claim 50; Fig 45; 286pp; English.

XX PS The present invention describes an isolated antibody that binds to one of the human PRO proteins designated PRO212, PRO290, PRO341, PRO535, PRO619, PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025, PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187, PRO1281, PRO1293, PRO139, PRO1317, PRO1710, PRO2094, PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The PRO polypeptides and nucleotides are useful in the treatment, diagnosis and prevention of cancer. The antibodies and other anti-tumour compounds may be used to treat various conditions, including those characterised by overexpression and/or activation of the amplified PRO genes. Exemplary conditions or disorders to be treated with such antibodies and other compounds include benign or malignant tumours (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and neck tumours), leukaemias and lymphoid malignancies, other disorders such as neuronal, glial,

CC astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal and blastocoele disorders, and inflammatory, angiogenic and immunologic disorders. AAC58242 to AAC58366 represent PCR primers and hybridisation probes used in the isolation of the human PRO sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO polynucleotide and protein sequences given in the exemplification of the present invention

SQ Sequence 908 BP; 173 A; 308 C; 310 G; 117 T; 0 U; 0 Other;

Query Match 0.9%; Score 20; DB 3; Length 908;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62

Db 254 GCAACAGCAGCAGCAGCGGC 273

RESULT 57

AAL49917

ID AAL49917 standard; DNA; 936 BP.

XX AC AAL49917;

XX DT 05-DEC-2002 (first entry)

XX DE Human pro-opiomelanocortin coding sequence fragment.

XX KW Human; pro-opiomelanocortin; POMC; corticosteroid; pain; analgesic; chronic inflammation; neuroendocrine; CRF; beta-endorphin; gene therapy; corticotrophin-releasing factor; antiinflammatory; gene, ds.

XX OS Homo sapiens.

XX PN WO200267996-A2.

XX PD 06-SEP-2002.

XX PF 19-FEB-2002; 2002WO-DE000583.

XX PR 24-FEB-2001; 2001DE-01009092.

XX PA (MOLO-) MOLOGEN FORSCH ENTWICKLUNGS & VERTRIEBS.

XX PI Wittig B, Stein C, Schaefer M, Schroff M, Junghans C; Koenig Merediz SA;

XX DR WPI; 2002-682796/73.

XX PT Agent for reducing pain, useful particularly for chronic inflammatory pain, comprises a nucleic acid construct for local expression of neuropeptides.

XX PS Claim 22; Page 33; 34pp; German.

XX CC The present invention relates to an agent for reducing or suppressing pain in higher animals, especially humans, which comprises (free from cells or cell lysate) an expression construct for local expression of a neuroendocrine peptide. The agent is especially useful for treating pain associated with chronic inflammatory disorders. The neuroendocrine may be pro-opiomelanocortin (POMC) or corticotrophin-releasing factor (CRF). The present sequence is a fragment of the human POMC coding sequence

SQ Sequence 936 BP; 190 A; 297 C; 328 G; 121 T; 0 U; 0 Other;

Query Match 0.9%; Score 20; DB 6; Length 936;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62

Db 401 GCAACAGCAGCAGCAGCGGC 420

```
RESULT 58
ABL27131
ID ABL27131 standard; DNA; 984 BP.
XX
AC ABL27131;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32866.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmacological; gene; ds.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 32866; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 984 BP; 332 A; 268 C; 225 G; 159 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 4; Length 984;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 38 CAGCGGCAACAGCAGCAGCA 57
DB 825 CAGCGGCAACAGCAGCAGCA 844
RESULT 59
ACN42014
ID ACN42014 standard; cDNA; 1038 BP.
XX
AC ACN42014;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:889.
XX
KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW dithp.
XX
OS Homo sapiens.
XX
FN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
XX
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthorne TA, Suchorski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patary S, Shi X, Suarez CJ;
XX
WPI; 2004-329368/30.
XX
P-PSDB; ABM83362.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 1; Page; 190pp; English.
XX
CC The invention relates to 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders, endocrine
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp polynucleotide of
CC the invention. Note: The sequence data for this patent is not represented
CC in the printed specification, but was obtained in electronic format
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 1038 BP; 208 A; 345 C; 355 G; 130 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 13; Length 1038;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GCAACAGCAGCAGCAGCGGC 62
DB 414 GCAACAGCAGCAGCAGCGGC 433
RESULT 60
ACN42013
ID ACN42013 standard; cDNA; 1039 BP.
XX
AC ACN42013;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:888.
XX
KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW dithp.
XX
```

```
OS Homo sapiens.
XX WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder IV;
XX PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
XX PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX PI Petalita CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
XX PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
XX PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX PI Patry S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
XX P-PSDB; ABM83361.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX PT in diagnosing a condition, disease or disorder associated with human
XX PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX PT in gene mapping.
XX PS Claim 1; Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
XX CC selected from one of the 2722 sequences defined in the specification. A
XX CC polynucleotide of the invention may have a use in gene therapy. The human
XX CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX CC used to diagnose a particular condition, disease or disorder associated
XX CC with human molecules, e.g. cell proliferative disorders,
XX CC autoimmune/inflammatory disorder, developmental disorder, endocrine
XX CC disorder, neurological disorders, gastrointestinal disorders, or
XX CC infections caused by virus, bacteria, fungi or parasite. The dithp
XX CC molecules may also be used in genetic mapping, in identifying individuals
XX CC from minute biological samples, in detecting single nucleotide
XX CC polymorphisms, as molecular weight markers, and for somatic or germline
XX CC gene therapy. The present sequence represents a dithp polynucleotide of
XX CC the invention. Note: The sequence data for this patent is not represented
XX CC in the printed specification, but was obtained in electronic format
XX CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 1039 BP; 218 A; 337 C; 347 G; 137 T; 0 U; 0 Other;
XX Query Match 0.9%; Score 20; DB 13; Length 1039;
XX Best Local Similarity 100.0%; Pred. No. 1.6e+02;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 43 GCAACAGCAGCAGCAGCGGC 62
DB 414 GCAACAGCAGCAGCAGCGGC 433
RESULT 61
ABL16775/c
ID ABL16775 standard; DNA; 1070 BP.
XX AC ABL16775;
XX XX 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 1798.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder IV;
OS Drosophila melanogaster.
XX WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 1798; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1070 BP; 291 A; 321 C; 277 G; 181 T; 0 U; 0 Other;
XX Query Match 0.9%; Score 20; DB 4; Length 1070;
XX Best Local Similarity 100.0%; Pred. No. 1.6e+02;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 37 GCAGCGGCAACAGCAGCAGC 56
DB 700 GCAGCGGCAACAGCAGCAGC 681
RESULT 62
ACN42012
ID ACN42012 standard; cDNA; 1088 BP.
XX AC ACN42012;
XX XX 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:887.
XX KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
XX KW dithp.
XX OS Homo sapiens.
XX WPI; 2004-329368/30.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder IV;
```



```

PR 19-MAR-1998; 98US-0078566P.
PR 19-MAR-1998; 98US-0078573P.
PR 19-MAR-1998; 98US-0078574P.
PR 19-MAR-1998; 98US-0078576P.
PR 19-MAR-1998; 98US-0078577P.
PR 19-MAR-1998; 98US-0078578P.
PR 19-MAR-1998; 98US-0078579P.
PR 19-MAR-1998; 98US-0078581P.
PR 01-APR-1998; 98US-0080312P.
PR 01-APR-1998; 98US-0080313P.
PR 01-APR-1998; 98US-0080314P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
PI Wei Y, Endress GA, Duan RD, Kyaw H, Ebner R, Lafleur DW, Olsen HS;
PI Shi Y, Moore PA;
XX
XX WPI; 1999-562050/47.
DR P-PSDB; AAY41377.
XX
XX New isolated human genes, useful for diagnosis and treatment of e.g.
PT cancers, neurological disorders, immune diseases, inflammation or blood
PT disorders.
XX
XX Claim 1; Page 340-341; 484pp; English.
XX
XX This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number, and the clone it is derived from, are
CC detailed in the descriptor line. The gene can be used to generate fusion
CC proteins by linking to the gene to a human immunoglobulin Fc portion
CC (e.g. AAZ24802) for increasing the stability of the fused protein as
CC compared to the human protein only. The invention relates to 95 novel
CC genes and their fragments (nucleic acid sequences: AAZ24811-Z24907; amino
CC acid sequences AAY41308-Y41404) which are useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
CC pathological conditions can be diagnosed by determining the amount of the
CC new polypeptides in a sample or by determining the presence of mutations
CC in the new polynucleotides. Specific uses are described for each of the
CC 95 polynucleotides, based on which tissues they are most highly expressed
CC in (see AAZ24811 for described uses)
XX
XX Sequence 1109 BP; 103 A; 454 C; 387 G; 165 T; 0 U; 0 Other;
SQ
Query Match 0.9%; Score 20; DB 2; Length 1109;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GCAACAGCAGCAGCAGCGGC 62
Db 152 GCAACAGCAGCAGCAGCGGC 133
RESULT 65
AAAL3712
ID AAA13712 standard; DNA; 1230 BP.
XX
XX AAA13712;
AC
XX
XX 25-JUL-2000 (first entry)
DT
XX
XX Human proopiomelanocortin gene exon 3 nucleotide sequence.
DE
XX
XX Human; proopiomelanocortin; POMC; beta-endorphin; opioid; addiction;
KW amphetamine; nervous system; ds.
KW
XX
XX Homo sapiens.
OS
XX
XX NO9804992-A.
PN
XX
XX 28-APR-2000.
FD
XX
XX 27-OCT-1998; 98NO-00004992.
PF

```

---

```

XX 27-OCT-1998; 98NO-00004992.
XX
XX (GENO-) GENOVA AS.
PA
XX
XX Fagerlund TH, Alestrom P, Berg K;
PI
XX
XX WPI; 2000-373908/32.
DR
XX
XX Novel isolated and purified nucleic acid molecule encoding beta-
PT endorphin, useful for treating a patient for psychological dependence on,
PT and abuse of, substances which have a stimulating and euphoric effect.
XX
XX Disclosure; Page 3-4; 8pp; Norwegian.
XX
XX The present invention describes recombinant plasmids for peripheral and
CC central nervous system analysing abuse of euphoric and stimulation
CC agents. The present sequence represents the nucleotide sequence of exon 3
CC from the human proopiomelanocortin (POMC) gene, which is used in the
CC exemplification of the present invention
XX
XX Sequence 1230 BP; 265 A; 379 C; 411 G; 175 T; 0 U; 0 Other;
SQ
Query Match 0.9%; Score 20; DB 3; Length 1230;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GCAACAGCAGCAGCAGCGGC 62
Db 474 GCAACAGCAGCAGCAGCGGC 493
RESULT 66
ADE28660/C
ID ADE28660 standard; cDNA; 1245 BP.
XX
XX ADE28660;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human NOV11a cDNA - SEQ ID 37.
DE
XX
XX NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW osteopathic; antiarthritic; antiinflammatory; dermatological;
KW antiasthmatic; antilipemic; metabolic; diabetes; obesity; infectious;
KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; human; ss; gene; NOV.
XX
XX Homo sapiens.
OS
XX
XX WO2003040330-A2.
PN
XX
XX 15-MAY-2003.
PD
XX
XX 05-NOV-2002; 2002WO-US035536.
PF
XX
XX 05-NOV-2001; 2001US-0338626P.
PR
XX
XX 05-DEC-2001; 2001US-0336600P.
PR
XX
XX 07-DEC-2001; 2001US-0338285P.
PR
XX
XX 12-DEC-2001; 2001US-0341346P.
PR
XX
XX 17-DEC-2001; 2001US-0341477P.
PR
XX
XX 17-DEC-2001; 2001US-0341540P.
PR
XX
XX 20-DEC-2001; 2001US-0342592P.
PR
XX
XX 27-DEC-2001; 2001US-0344297P.
PR
XX
XX 31-DEC-2001; 2001US-0344903P.
PR
XX
XX 17-APR-2002; 2002US-0373288P.
PR
XX
XX 15-MAY-2002; 2002US-0380981P.
PR

```

PR 17-MAY-2002; 2002US-0381495P.  
 PR 28-MAY-2002; 2002US-0383534P.  
 PR 28-MAY-2002; 2002US-0383744P.  
 PR 29-MAY-2002; 2002US-0383829P.  
 PR 29-MAY-2002; 2002US-0384024P.  
 PR 07-AUG-2002; 2002US-0401788P.  
 PR 26-AUG-2002; 2002US-0406353P.  
 PR 31-OCT-2002; 2002US-00287971.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL, Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A; Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L; Grose WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV; Lepley DM, Li L, MacDougall JR, Malyankar UM, Mazur A, McQueeney K; Mezes PS, Miller CE, Millet I, Mishra VS, Padigar M, Patturajan M; Pena CE, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA; Smithson G, Starling G, Szytek KA, Stone DJ, Tchernev VT, Twomlow N; Vernet CAM, Zerhusen BD, Zhong M;  
 DR WPI: 2003-441555/41.  
 DR F-PSDB; ADE28661.  
 XX  
 XX New isolated NOVX polypeptides and polynucleotides, useful for preventing, diagnosing or treating NOVX-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, asthma, or infections.  
 PT  
 PT Claim 20; SEQ ID NO 37; 447pp; English.  
 XX  
 XX The invention relates to a novel isolated NOVX polypeptide. The polypeptide of the invention demonstrates, antidiabetic, anorectic, radiant, hypotensive, antiarteriosclerotic, virucide, antibacterial, fungicide, protoscidic, nootropic, neuroprotective, antiparkinsonian, anticonvulsant, osteopathic, antiarthritic, antiinflammatory, dermatological, antiaesthetic and antilipemic activities. The polypeptides, nucleic acid molecules and antibodies may be useful for treating or diagnosing diseases including metabolic disorders such as diabetes and obesity, infectious diseases, anorexia, cancer, cardiovascular diseases including hypertension and atherosclerosis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic disorders, inflammatory skin disorders, asthma and dyslipidaemia. Furthermore, the nucleic acids and polypeptides may also be used to identify molecules that modulate or inhibit neurogenesis, cell differentiation and proliferation, haemopoiesis, wound healing and angiogenesis, as well as in gene therapy. Finally, the nucleic acids may be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The current sequence is that of the human NOVX cDNA of the invention.  
 CC  
 XX  
 SQ Sequence 1245 BP; 170 A; 486 C; 407 G; 182 T; 0 U; 0 Other;  
 Query Match 0.9%; Score 20; DB 10; Length 1245;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 43 GCACAGCAGCAGCAGCGC 62  
 DB 138 GCACAGCAGCAGCAGCGC 119  
 RESULT 67  
 ADM93405/c  
 ID ADM93405 standard; cDNA; 1245 BP.  
 XX  
 AC ADM93405;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human NOVX polynucleotide #19.  
 XX

KW Human; NOVX; gene; ss; congenital heart defect; cardiomyopathy; atherosclerosis; hypertension; pulmonary stenosis; scleroderma; adenocarcinoma; haemophilia; graft-versus-host disease; cancer; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; multiple sclerosis; diabetes; obesity; bronchial asthma; acquired immunodeficiency syndrome; AIDS; Crohn's disease; infectious disease; anorexia; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004067882-A1.  
 XX  
 PD 08-APR-2004.  
 XX  
 PF 05-NOV-2002; 2002US-00287971.  
 XX  
 PR 22-OCT-2001; 2001US-00035568.  
 PR 05-NOV-2001; 2001US-0386266P.  
 PR 06-NOV-2001; 2001US-033072P.  
 PR 09-NOV-2001; 2001US-0345398P.  
 PR 09-NOV-2001; 2001US-0348283P.  
 PR 15-NOV-2001; 2001US-0335610P.  
 PR 21-NOV-2001; 2001US-0332152P.  
 PR 28-NOV-2001; 2001US-0333912P.  
 PR 29-NOV-2001; 2001US-00997425.  
 PR 29-NOV-2001; 2001US-0334300P.  
 PR 04-DEC-2001; 2001US-033576P.  
 PR 05-FEB-2002; 2002US-0354807P.  
 PR 15-MAY-2002; 2002US-0380368P.  
 PR 16-MAY-2002; 2002US-0381043P.  
 PR 02-JUL-2002; 2002US-0393148P.  
 PR 02-JUL-2002; 2002US-0393262P.  
 PR 06-AUG-2002; 2002US-0401479P.  
 PR 06-AUG-2002; 2002US-0401626P.  
 PR 07-AUG-2002; 2002US-0401593P.  
 PR 07-AUG-2002; 2002US-0401695P.  
 PR 26-AUG-2002; 2002US-0406181P.  
 XX  
 XX (ALSO/) ALSOBROOK J P.  
 PA (ALVA/) ALVAREZ E.  
 PA (ANDE/) ANDERSON D W.  
 PA (BARO/) BARON M.  
 PA (BOLD/) BOLDOG F L.  
 PA (BURG/) BURGESS C E.  
 PA (CASW/) CASMAN S J.  
 PA (CHAP/) CHAPOVAL A.  
 PA (DHAN/) DHANABAL M.  
 PA (EDIN/) EDINGER S R.  
 PA (EISE/) EISEN A.  
 PA (ELLE/) ELLERMAN K.  
 PA (ETTE/) ETTEMBERG S.  
 PA (GANG/) GANGOLLI E A.  
 PA (GERL/) GERLACH V.  
 PA (GORM/) GORMAN L.  
 PA (GROS/) GROSSE W M.  
 PA (GUOX/) GUO X.  
 PA (HACK/) HACKETT C.  
 PA (JIW/) JI W.  
 PA (KEKU/) KEKUDA R.  
 PA (KRAH/) KHRANTSOV N V.  
 PA (LEPL/) LEPLEY D M.  
 PA (LILL/) LI L.  
 PA (MACD/) MACDOUGALL J R.  
 PA (MALY/) MALYANKAR U M.  
 PA (MAZU/) MAZUR A.  
 PA (MCQU/) MCQUEENEY K.  
 PA (MEZE/) MEZES P S.  
 PA (MILL/) MILLER C E.  
 PA (MILM/) MILLET I.  
 PA (MISH/) MISHRA V.  
 PA (PADI/) PADIGARU M.  
 PA (PATI/) PATTURAJAN M.  
 PA (PENA/) PENNA C E A.



PA (PEYM/) PEYMAN J A.  
PA (RST/) RASTELLI L.  
PA (RIEG/) RIEGER D K.  
PA (ROTH/) ROTHENBERG M E.  
PA (SHEN/) SHENOY S G.  
PA (SHIM/) SHIMKETS R A.  
PA (SMIT/) SMITHSON G.  
PA (SPAD/) SPADERNA S K.  
PA (STAR/) STARLING G.  
PA (SPYT/) SPYTEK K A.  
PA (STON/) STONE D J.  
PA (TCHE/) TCHERNEV V T.  
PA (TWO/) TWOMLOW N.  
PA (VERN/) VERNET C A M.  
PA (ZERH/) ZERHUSEN B D.  
PA (VOSS/) VOSS E Z.  
PA (ZHON/) ZHONG M.  
XX  
PI Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL, Burgeas CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A, Ellerman K, Etenberg S, Gangolli EA, Gerlach V, Gorman L, Grose WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV, Lepley DW, Li L, Macdougall JR, Malyankar UM, Mazur A, McQueeney K, Mezes PS, Miller CE, Millet I, Mishra V, Padigaru M, Patturajan M, Pena CE, Peyman JA, Rastelli L, Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G, Spaderna SK, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N, Vernet CAM, Zerhusen BD, Voss EZ, Zhong M;  
XX WPI: 2004-355303/33.  
DR P-PSDB; ADM93406.  
XX  
XX Novel isolated NOVX polypeptide useful treating or preventing disorders or syndromes such as Alzheimer's disease, Parkinson's disease, multiple sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.  
XX Claim 20; SEQ ID NO 37; 330bp; English.  
XX  
XX The invention relates to human NOVX polypeptides and the polynucleotides encoding them. The NOVX polypeptides and polynucleotides are useful for determining the presence of or predisposition to a disease associated with altered levels of the sequences in a mammalian subject, and for treating or preventing a pathology associated with NOVX. The polypeptides, polynucleotides and antibodies that bind immunospecifically to the polypeptides are useful for treating or preventing disorders or syndromes such as congenital heart defects, cardiomyopathy, atherosclerosis, hypertension, pulmonary stenosis, scleroderma, adenocarcinoma, haemophilia, graft-versus-host disease, cancer, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, multiple sclerosis, diabetes, obesity, bronchial asthma, acquired immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease, anorexia and immune disorders. This sequence represents a human NOVX polynucleotide of the invention. Note: The sequence data for this patent is also available from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
SQ Sequence 1245 BP; 170 A; 486 C; 407 G; 182 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 12; Length 1245;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 GCACAGCAGCAGCAGCGGC 62  
DB 138 GCACAGCAGCAGCAGCGGC 119  
|||||  
RESULT 68  
ACC79031/c  
ID ACC79031 standard; cDNA; 1418 BP.  
XX  
XX ACC79031;  
XX  
DT 30-JUL-2003 (first entry)

XX Human secreted protein SECP-6 encoding cDNA SEQ ID NO:86.  
XX Human; secreted protein; SECP; cytostatic; antiarteriosclerotic;  
KW anticonvulsant; nootropic; neuroprotective; cerebroprotective; anti-HIV;  
KW antiinflammatory; antiinflammatory; thymimetic; gene therapy; cancer;  
KW cell proliferative disorder; atherosclerosis; neurological disorder;  
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy; AIDS;  
KW inflammatory disorder; developmental disorder; hypothyroidism;  
KW Cushing's syndrome; infection; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO2003016506-A2.  
XX  
PD 27-FEB-2003.  
XX  
XX 15-AUG-2002; 2002WO-US027143.  
XX  
PR 17-AUG-2001; 2001US-0313249P.  
PR 24-AUG-2001; 2001US-0314752P.  
PR 07-SEP-2001; 2001US-0317818P.  
PR 07-SEP-2001; 2001US-0317824P.  
PR 21-SEP-2001; 2001US-0324040P.  
PR 24-SEP-2001; 2001US-0324586P.  
PR 02-NOV-2001; 2001US-0343980P.  
PR 28-NOV-2001; 2001US-0334229P.  
PR 13-FEB-2002; 2002US-0357002P.  
PR 06-MAR-2002; 2002US-0362439P.  
PR 19-MAR-2002; 2002US-0366041P.  
PR 30-APR-2002; 2002US-0376988P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Tang TY, Warren BA, Gietzen KJ, Lal PG, Yue H, Honchell CD, Lehr-Mason PM, Burford N, Xu Y, Baughn MR, Duggan BM, Tran UK, Lee EA, Forsythe IJ, Richardson TW, Lee S, Thangavelu K, Yue H, Emerling BM, Walla NK, Azimzai Y, Sanjanwala B, Hafalia AJA, Borowsky ML, Nguyen DB, Ison CH, Astronoff A, Ding L, Lee SY, Becha SD, Ramkumar J, Gandhi AR, Jin P, Fu GK, Swarnakar A;  
XX WPI: 2003-278569/27.  
DR P-PSDB; ABR56731.  
XX  
XX New human secreted proteins (SECP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant SECP expression e.g. cancer, AIDS, atherosclerosis, Huntington's disease, stroke, infections.  
XX  
PS Claim 5; Page 246; 286pp; English.  
XX  
XX ACC79026 to ACC79105 encode the human secreted proteins (I) given in ABR56726 to ABR56805, designated SECP-1 to SECP-80. SECP sequences can have cytostatic, antiarteriosclerotic, anticonvulsant, antiinflammatory, neuroprotective, cerebroprotective, anti-HIV, antiallergic and nootropic activities, and can be used in gene therapy. The SECP proteins and polynucleotides can be used in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or overexpression of SECP, such as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome) disorders, or infections. They are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of SECP. The SECP or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide  
XX  
SQ Sequence 1418 BP; 168 A; 542 C; 442 G; 266 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 10; Length 1418;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62  
| | | | | | | | | | | | | | | | | | | | | |  
Db 138 GCAACAGCAGCAGCAGCGGC 119

RESULT 69  
ABL04193  
ID ABL04193 standard; cDNA; 1565 BP.  
XX AC  
XX ABL04193;  
XX AC  
XX 26-MAR-2002 (first entry)  
XX XX  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 7061.  
XX DE  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX KW  
XX pharmaceutical; gene; ss.  
XX KW  
XX Drosophila melanogaster.  
XX OS  
XX WO200171042-A2.  
XX PN  
XX 27-SEP-2001.  
XX PD  
XX 23-MAR-2001; 2001WO-US009231.  
XX XX  
XX 23-MAR-2000; 2000US-0191637P.  
XX PR  
XX 11-JUL-2000; 2000US-00614150.  
XX PR  
XX (PEKE ) PE CORP NY.  
XX PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX PI  
XX WPI; 2001-65860/75.  
XX DR  
XX P-PSDB; ABB60090.  
XX DR  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX PT  
XX interactions.  
XX PT  
XX Claim 1; SEQ ID NO 7061; 21bp + Sequence Listing; English.  
XX PS  
XX The invention relates to an isolated nucleic acid detection reagent  
XX CC  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX CC  
XX useful in developmental biology and in elucidating cell signalling and  
XX CC  
XX cell-cell interactions in higher eukaryotes for the development of  
XX CC  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX CC  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX CC  
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
XX CC  
XX ABB72072). The sequence data for this patent did not form part of the  
XX CC  
XX printed specification, but was obtained in electronic format directly  
XX CC  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX CC  
XX Sequence 1565 BP; 397 A; 491 C; 416 G; 261 T; 0 U; 0 Other;  
XX SQ

Query Match 0.9%; Score 20; DB 4; Length 1565;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62  
| | | | | | | | | | | | | | | | | | | | | |  
Db 234 GCAACAGCAGCAGCAGCGGC 253

RESULT 70  
ADM47862  
ID ADM47862 standard; DNA; 1891 BP.  
XX AC  
XX ADM47862;  
XX AC  
XX 03-JUN-2004 (first entry)  
XX DT

XX DE  
XX XX  
XX Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;  
XX KW  
XX osmotic stress; sugar transport; cell cycle pathway; plant height;  
XX KW  
XX carbohydrate transport; crop productivity; plant growth;  
XX KW  
XX stress resistance; disease resistance; insect resistance; heat tolerance;  
XX KW  
XX nitrogen assimilation; water stress tolerance;  
XX KW  
XX photosynthetic carbon fixation; virus resistance; gene therapy; ds.  
XX XX  
XX Oryza sativa.  
XX AC  
XX US2003233670-A1.  
XX PN  
XX 18-DEC-2003.  
XX PD  
XX 04-DEC-2002; 2002US-00310154.  
XX PF  
XX 04-DEC-2001; 2001US-0337358P.  
XX PR  
XX (EDGE/) EDGERTON M D.  
XX PA  
XX (CHOM/) CHOMET P S.  
XX PA  
XX (LACC/) LACCETTI L B.  
XX XX  
XX Edgerton MD, Chomet PS, Laccetti LB;  
XX PI  
XX WPI; 2004-061374/06.  
XX DR  
XX P-PSDB; ADM48230.  
XX DR  
XX New polynucleotide, useful for manipulating plant protein quality,  
XX PT  
XX improving plant growth, yield and crop productivity or grain composition  
XX PT  
XX or producing plants with improved properties.  
XX PT  
XX Claim 1; SEQ ID NO 280; 144pp; English.  
XX PS  
XX The present invention relates to polynucleotide sequences, and the  
XX CC  
XX proteins they encode. The sequences are isolated from a variety of  
XX CC  
XX organisms such as plants (e.g. maize, rice, sorghum, thale cress,  
XX CC  
XX soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The  
XX CC  
XX polynucleotide and polypeptide sequences of the invention are useful in  
XX CC  
XX the production of transgenic plants that have improved properties. Also  
XX CC  
XX disclosed are methods of producing fertile transgenic plants, preferably  
XX CC  
XX maize, with desired phenotypes. The polynucleotide and polypeptide  
XX CC  
XX sequences are useful for improving plants by providing protection against  
XX CC  
XX osmotic stress, improving altering sugar transport and/or metabolism,  
XX CC  
XX modifying the cell cycle pathway, reducing plant height, modifying  
XX CC  
XX carbohydrate transport, improving crop productivity, improving plant  
XX CC  
XX growth and stress resistance, improving disease resistance, improving  
XX CC  
XX insect resistance, improving cold or heat tolerance, improving nitrogen  
XX CC  
XX assimilation, improving stalk strength, improving water stress tolerance,  
XX CC  
XX improving photosynthetic carbon fixation, improving biotic and abiotic  
XX CC  
XX stress resistance, improving resistance to oxidative stress, providing  
XX CC  
XX increased vigour, reducing senescence, and conferring virus resistance.  
XX CC  
XX The present sequence represents a polynucleotide sequence of the  
XX CC  
XX invention. Note: The sequence data for this patent is not provided in the  
XX CC  
XX printed specification but is obtained in electronic format from the USPTO  
XX CC  
XX website at seqdata.uspto.gov.  
XX SQ

Sequence 1891 BP; 379 A; 669 C; 494 G; 349 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 12; Length 1891;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GCGGCAACAGCAGCAGCAGC 59  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1125 GCGGCAACAGCAGCAGCAGC 1144

RESULT 71  
ADA71941  
ID ADA71941 standard; DNA; 2000 BP.  
XX DT



```
DE Human cDNA sequence SEQ ID NO:18501.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX PD 07-FEB-2001.
XX
XX PF 28-JUL-2000; 2000EP-00116126.
XX
XX PR 29-JUL-1999; 99JP-00248036.
XX
XX PR 27-AUG-1999; 99JP-00300253.
XX
XX PR 11-JAN-2000; 2000JP-00118776.
XX
XX PR 02-MAY-2000; 2000JP-00183767.
XX
XX PR 09-JUN-2000; 2000JP-00241899.
XX
XX PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX
XX PS Claim 8; SEQ ID NO 18501; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dr primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX SQ Sequence 2370 BP; 639 A; 459 C; 440 G; 832 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 4; Length 2370;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 972 GAACAAACAAACAAAGCTTAA 991
DB 1151 GAAACAAACAAACAAAGTTAA 1132
RESULT 74
AAS72135
ID AAS72135 standard; cDNA; 2370 BP.
XX
XX AAS72135;
XX
XX 13-FEB-2002 (first entry)
DT
```

```
XX DNA encoding novel human diagnostic protein #7939.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX
XX PR 23-AUG-2000; 2000US-00649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX
XX DR P-PSDB; ABG07948.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX PS Claim 1; SEQ ID NO 7939; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 2370 BP; 770 A; 396 C; 532 G; 672 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 5; Length 2370;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 790 ATGAGGAACAAACAGTATAT 809
DB 2040 ATGAGGAACAAACAGTATAT 2059
RESULT 75
ADG31190/c
ID ADG31190 standard; DNA; 2803 BP.
XX
XX ADG31190;
XX
XX 26-FEB-2004 (first entry)
DT
XX Novel mouse gene #15.
XX
```

XX novel protein; enzyme; blood clotting disorder; circulatory disorder;  
KW mouse; murine; gene; ds.  
XX  
OS Mus musculus.  
XX WO2003089644-A1.  
XX  
XX 30-OCT-2003.  
XX  
XX 18-APR-2003; 2003WO-JP004983.  
XX  
XX 19-APR-2002; 2002JP-00117794.  
XX 26-APR-2002; 2002JP-00126250.  
XX 30-APR-2002; 2002JP-00128657.  
XX 02-MAY-2002; 2002JP-00130702.  
XX 02-MAY-2002; 2002JP-00130871.  
XX 02-MAY-2002; 2002JP-00130918.  
XX 02-MAY-2002; 2002JP-00130925.  
XX 04-DEC-2002; 2002JP-00132326.  
XX 04-DEC-2002; 2002JP-00152381.  
XX 04-DEC-2002; 2002JP-00352559.  
XX 04-DEC-2002; 2002JP-00352694.  
XX 04-DEC-2002; 2002JP-00352786.  
XX 04-DEC-2002; 2002JP-00352808.  
XX  
XX (RIKE ) RIKEN KK.  
XX (DNAP-) DNAPFORM KK.  
XX (MITU ) MITSUBISHI CHEM CORP.  
XX  
XX Hayashizaki Y, Kamiya M, Kubodera H;  
XX  
XX WPI; 2004-034530/03.  
XX P-ESDB; ADG31214.  
XX  
XX Protease inhibitors, proteases, phosphodiesterases, transferases,  
XX isomerases and substances that affect their activity or expression useful  
XX for treating associated disorders.  
XX  
XX Claim 13; SEQ ID NO 15; 524pp; Japanese.  
XX  
XX The invention comprises the amino acid and coding sequences of proteins  
XX which have enzyme activity or protease, phosphodiesterase, transferase,  
XX or isomerase inhibitor activity. The DNA and protein sequences of the  
XX invention are useful for treating disorders associated with the proteins,  
XX such as blood clotting or circulatory disorders. The present DNA sequence  
XX represents a novel mouse gene of the invention.  
XX  
XX Sequence 2803 BP; 546 A; 848 C; 854 G; 555 T; 0 U; 0 Other;  
SQ  
Query Match 0.9%; Score 20; DB 12; Length 2803;  
Best Local Similarity 100.0%; Pred.No. 1.5e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 43 GCAACAGCAGCAGCAGCGGC 62  
Db 338 GCAACAGCAGCAGCAGCGGC 319

Search completed: March 25, 2005, 17:53:06  
Job time : 1171 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 20:35:20 ; Search time 1231 Seconds  
(without alignments)  
11250.599 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggacgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 552208 seqs, 2979665951 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/ECT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/ECTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	9	US-09-903-248-3
2	2324	100.0	2324	9	US-09-859-604-3
3	2324	100.0	2324	9	US-09-903-063-3
4	2324	100.0	2324	9	US-09-903-216-3
5	2324	100.0	2324	9	US-09-903-199-3
6	2324	100.0	2324	9	US-09-903-023-3
7	2324	100.0	2324	10	US-09-436-184-3
8	2171	93.4	5358	15	US-10-084-817-63
9	895	38.5	2680	16	US-10-096-534-37
10	895	38.5	2680	16	US-10-263-909-3
11	640	27.5	2442	9	US-09-736-457-1791

12	640	27.5	2442	9	US-09-902-941-1791	Sequence 1791, Ap
13	640	27.5	2442	9	US-09-849-628-1791	Sequence 1791, Ap
14	640	27.5	2442	14	US-10-017-754-1791	Sequence 1791, Ap
15	640	27.5	2442	16	US-10-113-872-1791	Sequence 1791, Ap
16	640	27.5	2442	17	US-10-283-017-1791	Sequence 9938, Ap
17	615	26.5	3110	14	US-10-198-846-9938	Sequence 34279, A
18	459	19.8	502	10	US-09-918-995-34279	Sequence 3202, Ap
19	359	15.4	660	11	US-09-969-034-3202	Sequence 25661, A
20	315	13.6	366	17	US-10-242-535A-25661	Sequence 25661, A
21	315	13.6	366	17	US-10-085-783A-25661	Sequence 2707, Ap
22	249	10.7	495	14	US-10-066-543-277	Sequence 2707, Ap
23	135	5.8	191	17	US-10-242-535A-2707	Sequence 2707, Ap
24	135	5.8	191	17	US-10-085-783A-2707	Sequence 26964, A
25	104	4.5	161	16	US-10-029-386-26964	Sequence 13264, A
26	104	4.5	592	13	US-10-029-386-13264	Sequence 2825, Ap
27	78	3.4	450	13	US-10-027-632-2825	Sequence 2825, Ap
28	78	3.4	450	17	US-10-027-632-2825	Sequence 24, Appl
29	72	3.1	1382	17	US-10-443-622-24	Sequence 11070, A
30	60	2.6	60	10	US-09-908-975-11070	Sequence 322, Ap
31	59	2.5	2208	17	US-10-116-275-322	Sequence 177, App
32	59	2.5	2208	18	US-10-755-889-177	Sequence 3266, Ap
33	51	2.2	583	11	US-09-969-034-3266	Sequence 7020, Ap
34	44	1.9	591	13	US-10-027-632-7020	Sequence 7020, Ap
35	44	1.9	591	17	US-10-027-632-7020	Sequence 56205, A
36	39	1.7	453	13	US-10-027-632-56205	Sequence 310460, A
37	39	1.7	453	17	US-10-027-632-56205	Sequence 56205, A
38	39	1.7	453	17	US-10-027-632-56205	Sequence 310460, A
39	39	1.7	453	17	US-10-027-632-56205	Sequence 310460, A
40	24	1.0	859	17	US-10-424-559-97119	Sequence 97119, A
41	23	1.0	2135	18	US-10-425-115-124491	Sequence 124491, A
42	21	0.9	681	18	US-10-437-963-73067	Sequence 73067, A
43	21	0.9	681	18	US-10-425-115-42949	Sequence 42949, A
44	21	0.9	1845	16	US-10-308-503-188	Sequence 188, App
45	20	0.9	20	9	US-09-859-604-10	Sequence 10, Appl
46	20	0.9	20	9	US-09-859-604-11	Sequence 12, Appl
47	20	0.9	20	9	US-09-859-604-12	Sequence 21448, A
48	20	0.9	183	18	US-10-425-115-21348	Sequence 22449, A
49	20	0.9	489	18	US-10-767-701-26002	Sequence 26002, A
50	20	0.9	495	17	US-10-424-599-89196	Sequence 89196, A
51	20	0.9	550	16	US-10-029-386-9144	Sequence 9144, Ap
52	20	0.9	567	13	US-10-027-632-82527	Sequence 82527, A
53	20	0.9	567	13	US-10-027-632-322022	Sequence 322022, A
54	20	0.9	567	13	US-10-027-632-322023	Sequence 82527, A
55	20	0.9	567	17	US-10-027-632-82527	Sequence 322022, A
56	20	0.9	567	17	US-10-027-632-322023	Sequence 322023, A
57	20	0.9	750	18	US-10-425-115-112661	Sequence 112661, A
58	20	0.9	771	18	US-10-437-963-71261	Sequence 71261, A
59	20	0.9	804	18	US-10-822-613-1	Sequence 1, Appli
60	20	0.9	843	18	US-10-437-963-24406	Sequence 24406, A
61	20	0.9	924	18	US-10-437-963-51503	Sequence 51503, A
62	20	0.9	936	18	US-10-646-620A-8	Sequence 8, Appli
63	20	0.9	936	18	US-10-425-115-180542	Sequence 180542, A
64	20	0.9	1011	18	US-09-853-386-105	Sequence 105, App
65	20	0.9	1072	9	US-09-397-945-80	Sequence 80, Appl
66	20	0.9	1109	17	US-10-653-595-80	Sequence 80, Appl
67	20	0.9	1109	17	US-09-960-288-1	Sequence 1, Appli
68	20	0.9	1230	17	US-10-287-971-37	Sequence 37, Appl
69	20	0.9	1245	17	US-10-425-114-23644	Sequence 23644, A
70	20	0.9	1282	17	US-10-425-115-19062	Sequence 19062, A
71	20	0.9	1435	18	US-10-913-196-2	Sequence 2, Appli
72	20	0.9	1439	19	US-10-425-114-34740	Sequence 34740, A
73	20	0.9	1850	17	US-10-310-154-280	Sequence 280, App
74	20	0.9	1891	17	US-10-425-115-16475	Sequence 16475, A
75	20	0.9	1895	18	US-10-260-238-2021	Sequence 2021, Ap
76	20	0.9	2000	17	US-09-788-711A-3	Sequence 3, Appli
77	20	0.9	8772	9	US-09-788-711A-3	Sequence 1, Appli
78	20	0.9	8871	9	US-10-311-623-21	Sequence 21, Appl
79	20	0.9	9321	15	US-10-225-567A-523	Sequence 523, App
80	20	0.9	10531	15	US-10-292-798-931	Sequence 931, App
81	20	0.9	24370	17	US-10-741-600-17649	Sequence 17649, A
82	20	0.9	47777	19	US-10-087-132-130	Sequence 130, App
83	20	0.9	52677	13		

c 85 20 0.9 159400 18 US-10-450-826-33 Sequence 33, Appl  
86 20 0.9 822900 17 US-10-292-798-1393 Sequence 1393, Ap  
c 87 20 0.9 2731748 18 US-10-297-465A-1 Sequence 1, Appli  
c 88 19 0.8 193 18 US-10-437-963-88593 Sequence 88593, A  
c 89 19 0.8 233 17 US-10-424-599-1798 Sequence 1798, Ap  
c 90 19 0.8 347 18 US-10-425-115-14296 Sequence 14296, A  
c 91 19 0.8 353 9 US-09-783-590-4838 Sequence 4838, Ap  
c 92 19 0.8 389 9 US-09-960-352-868 Sequence 868, App  
c 93 19 0.8 394 9 US-09-960-352-3447 Sequence 3447, Ap  
c 94 19 0.8 400 9 US-09-960-352-7474 Sequence 7474, Ap  
c 95 19 0.8 404 9 US-09-960-352-8848 Sequence 8848, Ap  
96 19 0.8 483 13 US-10-062-254-197 Sequence 197, App  
97 19 0.8 497 17 US-10-425-114-29177 Sequence 29177, A  
c 98 19 0.8 500 17 US-10-621-901-185 Sequence 185, App  
c 99 19 0.8 505 18 US-10-437-963-85957 Sequence 85957, A  
c 100 19 0.8 507 10 US-09-918-995-15209 Sequence 15209, A

## ALIGNMENTS

## RESULT 1

US-09-903-248-3  
; Sequence 3, Application US/09903248  
; Patent No. US20020102263A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV5  
; CURRENT APPLICATION NUMBER: US/09/903,248  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-248-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCATGGCCAGCGTGAAGATGCCAGAGCAGCGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCATGGCCAGCGTGAAGATGCCAGAGCAGCGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGCGAGGTAGCAGAGTGGGCGAGCAGCAGCGCCCGGCGCGGAGAGACAA 120  
Db 61 GCTCCGCGAGGTAGCAGAGTGGGCGAGCAGCAGCGCCCGGCGCGGAGAGACAA 120  
Qy 121 AGCATGAGGACACAAGAAATGGGAGGAAAGCGGACTCTCGGAACTTCATTCTTCAGT 180  
Db 121 AGCATGAGGACACAAGAAATGGGAGGAAAGCGGACTCTCGGAACTTCATTCTTCAGT 180  
Qy 181 GGTATTAGTGCATTCGATTCGTGGCGCTCGGACATCTGTAGCTGTGCTTTGGTTGATC 240  
Db 181 GGTATTAGTGCATTCGATTCGTGGCGCTCGGACATCTGTAGCTGTGCTTTGGTTGATC 240  
Qy 241 TTGTTGACTATGAGGAGTCTTAGGAAATCTAGAACTATGATGCTGATGCTGATGGAG 300  
Db 241 TTGTTGACTATGAGGAGTCTTAGGAAATCTAGAACTATGATGCTGATGCTGATGGAG 300  
Qy 301 ATTTTGTGATGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTTCAGAGC 360  
Db 301 ATTTTGTGATGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCCAGAGAGCGCTGAGCCACACACTGAGCCCGCAGGAGCGTTCTCTGTGG 420

Db 361 CAGCAGTCCCGCCAGAGAGCGCTGAGCCACACACTGAGCCCGGAGGAGCAGTTCCTGTGG 420  
Qy 421 AGGCAGAAACCCAGAAATATCGAAGATGAAGCAAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGCAGAAACCCAGAAATATCGAAGATGAAGCAAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGCGAGAACTGTTGAGGAGAAAGCTTTCACAAAGAAAGATGGACCCACAG 540  
Db 481 AAATGGTACACGCGAGAACTGTTGAGGAGAAAGCTTTCACAAAGAAAGATGGACCCACAG 540  
Qy 541 GAGAACCAACCAAGAGGATGATGAGTTCTTATGGCGACTGATGTAGATGATGATTTG 600  
Db 541 GAGAACCAACCAAGAGGATGATGAGTTCTTATGGCGACTGATGTAGATGATGATTTG 600  
Qy 601 AGACCCCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGGAAGAGA 660  
Db 601 AGACCCCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGGAAGAGA 660  
Qy 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTCGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTCGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACCAAGTATGGAAGATGAAGATTTGCAACCATGATACAGATGATGAACAT 780  
Db 721 ATTCCAGTGAACCAAGTATGGAAGATGAAGATTTGCAACCATGATACAGATGATGAACAT 780  
Qy 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGAGATGAATCA 840  
Db 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGAGATGAATCA 840  
Qy 841 CAGAAGTAACTGCTCCCGCTGAGGATTAATCCTGTAGAAAGATTACAGGTAATTTAGAAG 900  
Db 841 CAGAAGTAACTGCTCCCGCTGAGGATTAATCCTGTAGAAAGATTACAGGTAATTTAGAAG 900  
Qy 901 AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAGTACCACAGAAACAATAGAAAAA 960  
Db 901 AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAGTACCACAGAAACAATAGAAAAA 960  
Qy 961 CAGATGATCCAGAAACAAAAAGTAAAGAAAAAGAGCCCTTAAATCTTTAAATAAT 1020  
Db 961 CAGATGATCCAGAAACAAAAAGTAAAGAAAAAGAGCCCTTAAATCTTTAAATAAT 1020  
Qy 1021 TTGATAAGACTATTAAAGCTGAATTTGATGCTGAGAAAAAATCCGTTAAAGGGGAAAAA 1080  
Db 1021 TTGATAAGACTATTAAAGCTGAATTTGATGCTGAGAAAAAATCCGTTAAAGGGGAAAAA 1080  
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCCTCAGAGTCCAGAG 1140  
Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCCTCAGAGTCCAGAG 1140  
Qy 1141 CAAGATATGGAAAGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGAGTAATGAGG 1200  
Db 1141 CAAGATATGGAAAGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGAGTAATGAGG 1200  
Qy 1201 TGCTACGTGGAGCCATCGAGACCTTACCAAGAGTGGCCAGCCTACCTGATGTCCTGAG 1260  
Db 1201 TGCTACGTGGAGCCATCGAGACCTTACCAAGAGTGGCCAGCCTACCTGATGTCCTGAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTGAAGCGTCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTGAAGCGTCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
Qy 1321 GAGGTTCCCTGCTTACCTCGAGAGTATGTTCAACTATTTCCTCAATGATCTCTTAA 1380  
Db 1321 GAGGTTCCCTGCTTACCTCGAGAGTATGTTCAACTATTTCCTCAATGATCTCTTAA 1380  
Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGATAAATGACAAATGCAAGAAAGTTT 1440  
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGATAAATGACAAATGCAAGAAAGTTT 1440  
Qy 1441 ATGAGAGGTGCTGAGTGTGACACCTAATGATGGCTTGTGCTAAAGTCCATATGCTTCA 1500

Db 1441 ATGAGAGGTCTGAGTGTGACACCTAATGATGGCTTTGCTTAAAGTCCATATGCGTTCA 1500  
Qy TCTGAGGCACAGAACAAATTCGTGAGAGCATCCCATATTTAAAGAGGAATGAAT 1560  
Db TCTGAGGCACAGAACAAATTCGTGAGAGCATCCCATATTTAAAGAGGAATGAAT 1560  
Qy CCGAGATCTCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCGA 1620  
Db CCGAGATCTCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCGA 1620  
Qy GGTGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCACAGAGAGGACATTTG 1680  
Db GGTGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCACAGAGAGGACATTTG 1680  
Qy CATCTGCTGGCAACGCTCTCTACATGTGAATGGACTGAAGGACAGGCTTGTGGA 1740  
Db CATCTGCTGGCAACGCTCTCTACATGTGAATGGACTGAAGGACAGGCTTGTGGA 1740  
Qy CCCGAAAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACTGGAAGTTAA 1800  
Db CCCGAAAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACTGGAAGTTAA 1800  
Qy TCCGAGATGAAGGCTTTGCAGTGTATGATGATAAAGCCAAAGGCTCTTCTGCTGAGGATG 1860  
Db TCCGAGATGAAGGCTTTGCAGTGTATGATGATAAAGCCAAAGGCTCTTCTGCTGAGGATG 1860  
Qy AAAACCTGAGGAAAGAGGCTTGGAGCCAGTTCACGCTGTGGCAGCAGAGGAAGAA 1920  
Db AAAACCTGAGGAAAGAGGCTTGGAGCCAGTTCACGCTGTGGCAGCAGAGGAAGAA 1920  
Qy ATGAAATGCTGCAAGAGGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
Db ATGAAATGCTGCAAGAGGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
Qy CAACAGGATGAGAGGAGGACAGATCAAAATATTCATCATGACCCCGGAGCTCACGTGT 2040  
Db CAACAGGATGAGAGGAGGACAGATCAAAATATTCATCATGACCCCGGAGCTCACGTGT 2040  
Qy GCGCGCACACAGGCGCCACAACTCGAGGCTCCGAATGACCTGCGGCTTGGTATTCCTCA 2100  
Db GCGCGCACACAGGCGCCACAACTCGAGGCTCCGAATGACCTGCGGCTTGGTATTCCTCA 2100  
Qy AGAAGGCTGACAGATTCGATGTCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160  
Db AGAAGGCTGACAGATTCGATGTCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160  
Qy TCATCTTTGATGACTCTTTGAGCAGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220  
Db TCATCTTTGATGACTCTTTGAGCAGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220  
Qy TATTATCGTGGATGTTGGCATCCGGAATGACACACAGAGAGGAGGAGGAGGAGGAGG 2280  
Db TATTATCGTGGATGTTGGCATCCGGAATGACACACAGAGAGGAGGAGGAGGAGGAGG 2280  
Qy CAATTTAGCATGAATTCATGCAAGCTTTGGAACTCTGGAGAGA 2324  
Db CAATTTAGCATGAATTCATGCAAGCTTTGGAACTCTGGAGAGA 2324

## RESULT 2

US-09-859-604-3  
; Sequence 3, Application US/09859604  
; Patent No. US20020110559A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M  
; APPLICANT: Deutch, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/436,184

; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-859-604-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGCGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAA 120  
Db 61 GCTCCGCGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAA 120  
Qy 121 AGCATGGAGGACACAAAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
Db 121 AGCATGGAGGACACAAAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
Qy 181 GGTATTATGTTGATGATGATGCTGGCGCTCTGGAACATCTGTAGCTGTCTGTTGTTGATC 240  
Db 181 GGTATTATGTTGATGATGATGCTGGCGCTCTGGAACATCTGTAGCTGTCTGTTGTTGATC 240  
Qy 241 TTGTTGATATGAGGAGGAGTCTAGGAACTATAGGAACTATAGGAACTATAGGAACTATAG 300  
Db 241 TTGTTGATATGAGGAGGAGTCTAGGAACTATAGGAACTATAGGAACTATAGGAACTATAG 300  
Qy 301 ATTTTGTATGTTGATGATGCTGCAAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAG 360  
Db 301 ATTTTGTATGTTGATGATGCTGCAAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAG 360  
Qy 361 CAGCAGTCCCGCAGAGGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGAGGAGGAGG 420  
Db 361 CAGCAGTCCCGCAGAGGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGAGGAGGAGG 420  
Qy 421 AGGAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGACAGAACATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
Db 481 AAATGGTACACGACAGAACATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
Qy 541 GAGAAACCAACAAAGAGGATGATGATTTCTTATGGGAGTGTATGATGATGATGATGATGAT 600  
Db 541 GAGAAACCAACAAAGAGGATGATGATTTCTTATGGGAGTGTATGATGATGATGATGATGAT 600  
Qy 601 AGACCTGGGAACTCTGAGTATCTCATGAGAAACCGAGCATAGTTACCACTGGGAGGAGG 660  
Db 601 AGACCTGGGAACTCTGAGTATCTCATGAGAAACCGAGCATAGTTACCACTGGGAGGAGG 660  
Qy 661 CAGTTTCAAGAGCTGTAATCAGGATATGGAAGATGATGATGATGATGATGATGATGATGAT 720  
Db 661 CAGTTTCAAGAGCTGTAATCAGGATATGGAAGATGATGATGATGATGATGATGATGATGAT 720  
Qy 721 ATTCCAGTGAACCCAGTAGTAGAAGATGAAGATTTGCAACATGATACAGATGATGATGAT 780  
Db 721 ATTCCAGTGAACCCAGTAGTAGAAGATGAAGATTTGCAACATGATACAGATGATGATGAT 780  
Qy 781 ACCAAGTCTATGAGGAGCAACAGCATATATGAACTCTAGAAATGAGGAGGATGAGAACTCA 840  
Db 781 ACCAAGTCTATGAGGAGCAACAGCATATATGAACTCTAGAAATGAGGAGGATGAGAACTCA 840  
Qy 841 CAGAAGTAACTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 900  
Db 841 CAGAAGTAACTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 900



QY 901 AAGTAAGCATTTTTCCTGTCGAGAAACAGCAGGAGTACACACAGAAACAAATAGAAAA 960  
DB 901 AAGTAAGCATTTTTCCTGTCGAGAAACAGCAGGAGTACACACAGAAACAAATAGAAAA 960  
QY 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAAATCTTTTAAATAAT 1020  
DB 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAAATCTTTTAAATAAT 1020  
QY 1021 TTGATAAGCATTTTAAAGCTGAATCTGATGTCGAGAAACCTCCGTAAGAGGAAAAA 1080  
DB 1021 TTGATAAGCATTTTAAAGCTGAATCTGATGTCGAGAAACCTCCGTAAGAGGAAAAA 1080  
QY 1081 TTGAGGAGCAGTGAATGCTTTAAAGCACTAGTACGCAATACCTCCAGAGTCCACGAG 1140  
DB 1081 TTGAGGAGCAGTGAATGCTTTAAAGCACTAGTACGCAATACCTCCAGAGTCCACGAG 1140  
QY 1141 CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGGAATGATGAGG 1200  
DB 1141 CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGGAATGATGAGG 1200  
QY 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCGAG 1260  
DB 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCGAG 1260  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
DB 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
QY 1321 GAGGTTCCCTGCTTACCTCGAGAGATGATTTCACTAATTTCCCAATGATCTTCTTAA 1380  
DB 1321 GAGGTTCCCTGCTTACCTCGAGAGATGATTTCACTAATTTCCCAATGATCTTCTTAA 1380  
QY 1381 AAAATGACCTTGGCTGGGATACCTCTGTAGAGAGATAATGACAATGCAAAAGAAAGTTT 1440  
DB 1381 AAAATGACCTTGGCTGGGATACCTCTGTAGAGAGATAATGACAATGCAAAAGAAAGTTT 1440  
QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
DB 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
QY 1501 TCTGTAAGGCACAGAAACAAATGCTGAGAGCATCCATTTTAAAGGAAGGAATAGAT 1560  
DB 1501 TCTGTAAGGCACAGAAACAAATGCTGAGAGCATCCATTTTAAAGGAAGGAATAGAT 1560  
QY 1561 CCGAGATCTCGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGA 1620  
DB 1561 CCGAGATCTCGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGA 1620  
QY 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTGATGAGCTTGGGCAAGAGAGGACACTTTG 1680  
DB 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTGATGAGCTTGGGCAAGAGAGGACACTTTG 1680  
QY 1681 CATCTGTCTGGCAACGCTCACTCACTCAATGTAAGTGAAGTGAAGCAAGAGGAGGAGTGA 1740  
DB 1681 CATCTGTCTGGCAACGCTCACTCACTCAATGTAAGTGAAGTGAAGCAAGAGGAGGAGTGA 1740  
QY 1741 CCCCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAATCTGAAAGTTAA 1800  
DB 1741 CCCCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAATCTGAAAGTTAA 1800  
QY 1801 TCGAGATGAAGCCCTTGCAGTGTGATGATGAAGCCCAAGAGTCTCTCTGCTGAGGATG 1860  
DB 1801 TCGAGATGAAGCCCTTGCAGTGTGATGATGAAGCCCAAGAGTCTCTCTGCTGAGGATG 1860  
QY 1861 AAAACCTGAGGAAAGAGGAGCTGAGGAGTTCAGCTGTGGCAGCAAGGAAAGAA 1920  
DB 1861 AAAACCTGAGGAAAGAGGAGCTGAGGAGTTCAGCTGTGGCAGCAAGGAAAGAA 1920  
QY 1921 ATGAAATGCTCAGAAAGAGGCTCTTAAACCTGTACTTAAAGAAAGTTTCCCGAGA 1980  
DB 1921 ATGAAATGCTCAGAAAGAGGCTCTTAAACCTGTACTTAAAGAAAGTTTCCCGAGA 1980  
QY 1981 CAACAGGATGCAGAGAGGACAGATCAATATTTCCATCATGACCCCGGACTCAAGTGT 2040

DB 1981 CAACAGGATGCAGAGAGGACAGATCAATATTTCCATCATGACCCCGGACTCAAGTGT 2040  
QY 2041 GCGCGCACACAGGCGCCCAAAATGCGAGCTCCGAATGCACTGGGCTTGGTGAATTTCCCA 2100  
DB 2041 GCGCGCACACAGGCGCCCAAAATGCGAGCTCCGAATGCACTGGGCTTGGTGAATTTCCCA 2100  
QY 2101 AGGAAGCTTGCAGATTTCCGATGTGCCAACGAGACAGGACCTGGGAGGAAGCAAGGTGC 2160  
DB 2101 AGGAAGCTTGCAGATTTCCGATGTGCCAACGAGACAGGACCTGGGAGGAAGCAAGGTGC 2160  
QY 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGTATGCGAGATGCTTCTTCCGCTGA 2220  
DB 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGTATGCGAGATGCTTCTTCCGCTGA 2220  
QY 2221 TATTCATCTGTGATGTGTGCAATCCGAACTGACACACAGCAGAGACGACGCTTCCAG 2280  
DB 2221 TATTCATCTGTGATGTGTGCAATCCGAACTGACACACAGCAGAGACGACGCTTCCAG 2280  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
DB 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 3  
US-09-903-063-3  
; Sequence 3, Application US/09903063  
; Patent No. US20020114810A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R., Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV3  
; CURRENT APPLICATION NUMBER: US/09/903,063  
; CURRENT FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-903-063-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACCTGTCAATGGCCAGCGTAAGATGCCAAGACAGCGGCAACAGCAGCAGCAGCG 60  
DB 1 CCGACCTGTCAATGGCCAGCGTAAGATGCCAAGACAGCGGCAACAGCAGCAGCAGCG 60  
QY 61 GCTCCGCGCAGGTAGCACCGAGTGCAGGACAGCAGCCCGGGGCCCGGAGAGAGACAA 120  
DB 61 GCTCCGCGCAGGTAGCACCGAGTGCAGGACAGCAGCCCGGGGCCCGGAGAGAGACAA 120  
QY 121 AGCATGGAGGACACAAGAATGGGAGGAAGCGGACTCTCGGGAATCTTCACTTTCACGT 180  
DB 121 AGCATGGAGGACACAAGAATGGGAGGAAGCGGACTCTCGGGAATCTTCACTTTCACGT 180  
QY 181 GGTTCATGTCGATTTGCAATGCTGGCGCTCTGGACATCTGATGCTGCTTTGGTTTGCATC 240  
DB 181 GGTTCATGTCGATTTGCAATGCTGGCGCTCTGGACATCTGATGCTGCTTTGGTTTGCATC 240  
QY 241 TTGTTGACTATGAGCAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGATGAG 300  
DB 241 TTGTTGACTATGAGCAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGATGAG 300  
QY 301 ATTTTGTGTCGATGATGCCAAAGTTTATTTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360

Db 301 ATTTTCATGTGATGATGCCAAAGTTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGAGGTTCTCTGTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGAGGTTCTCTGTGG 420  
Qy 421 AGCAGAACCCCGAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCTCATG 480  
Db 421 AGCAGAACCCCGAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCTCATG 480  
Qy 481 AAATGTTTACAGCAGAAACATGTTGAGGGAGAAAGTTCGAAACAAAGAGATGGAACCCACAG 540  
Db 481 AAATGTTTACAGCAGAAACATGTTGAGGGAGAAAGTTCGAAACAAAGAGATGGAACCCACAG 540  
Qy 541 GAGAACCAACAAAGAGAGATGATGAGTTCCTTATGCGGACTGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACAAAGAGAGATGATGAGTTCCTTATGCGGACTGATGATGATGATGATGATG 600  
Qy 601 AGACCTCGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTTACCAAGTGAAGAGA 660  
Db 601 AGACCTCGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTTACCAAGTGAAGAGA 660  
Qy 661 CAGTTTCAAGAGATGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGAGATGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACAGTATGTAAGATGAAGATTTGACCATGATACAGATGATGTAACAT 780  
Db 721 ATTCCAGTGAACAGTATGTAAGATGAAGATTTGACCATGATACAGATGATGTAACAT 780  
Qy 781 ACCAAGTCTATGAGGAACAGCAGTATGTAAGATTTGAACTTGAAGATGAAGATGAATCA 840  
Db 781 ACCAAGTCTATGAGGAACAGCAGTATGTAAGATTTGAACTTGAAGATGAAGATGAATCA 840  
Qy 841 CAGAAGTAACTGCTCCCTCGAGGATTAATCCTGTGAAGATTTCAAGGTAATTTGTAGAAG 900  
Db 841 CAGAAGTAACTGCTCCCTCGAGGATTAATCCTGTGAAGATTTCAAGGTAATTTGTAGAAG 900  
Qy 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGATGACCCAGAAACAAATAGAAAAA 960  
Db 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGATGACCCAGAAACAAATAGAAAAA 960  
Qy 961 CAGATGATCCAGAAACAAAGATTTAAGAAAGAAAGCCTTAACTTTTAAATTAAT 1020  
Db 961 CAGATGATCCAGAAACAAAGATTTAAGAAAGAAAGCCTTAACTTTTAAATTAAT 1020  
Qy 1021 TTGATAAGCATTTAAGCTGAACCTTGATGCTGCAAGAAACCTCGTAAAGGGGAAAAA 1080  
Db 1021 TTGATAAGCATTTAAGCTGAACCTTGATGCTGCAAGAAACCTCGTAAAGGGGAAAAA 1080  
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAG 1140  
Qy 1141 CAAAGATGGAAGCGCAGTGTGAGGATGATTTGGCTGGAAGAGAGAGATTAATGAGG 1200  
Db 1141 CAAAGATGGAAGCGCAGTGTGAGGATGATTTGGCTGGAAGAGAGAGATTAATGAGG 1200  
Qy 1201 TGCTACGTGGAGCCATCGACCTACCAAGAGTGGCCAGCTACCTGATGCTCCCTGAG 1260  
Db 1201 TGCTACGTGGAGCCATCGACCTACCAAGAGTGGCCAGCTACCTGATGCTCCCTGAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCTGCTGAGCAAGAAACCTCGTAAAGGGGAAAAA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCTGCTGAGCAAGAAACCTCGTAAAGGGGAAAAA 1320  
Qy 1321 GAGGTTCCCTGTTACCTCGAGAGATTAAGTTCAACTATTTCCCAATGATCTCTCTTAA 1380  
Db 1321 GAGGTTCCCTGTTACCTCGAGAGATTAAGTTCAACTATTTCCCAATGATCTCTCTTAA 1380  
Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATTAATGCAATGCAAGAAAGATTT 1440  
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATTAATGCAATGCAAGAAAGATTT 1440

Qy 1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTAATGGCTTCA 1500  
Db 1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTAATGGCTTCA 1500  
Qy 1501 TCTTGAAGGCAACAGAAACAAATTTGCTGAGAGCATCCCATATTTTAAAGAAAGGAATAGAAT 1560  
Db 1501 TCTTGAAGGCAACAGAAACAAATTTGCTGAGAGCATCCCATATTTTAAAGAAAGGAATAGAAT 1560  
Qy 1561 CCGGAGATCTGGCAGCTGATGATGGGAGATTTTATTTTCCACTGGGGATGCCATGCAGA 1620  
Db 1561 CCGGAGATCTGGCAGCTGATGATGGGAGATTTTATTTTCCACTGGGGATGCCATGCAGA 1620  
Qy 1621 GGGTTGGGAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCAACAGAGAGGACACTTTG 1680  
Db 1621 GGGTTGGGAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCAACAGAGAGGACACTTTG 1680  
Qy 1681 CATCTGTCTGCAACCGCTCACTCAATGTAATGGAATGGAATGGAATGGAATGGAATGGA 1740  
Db 1681 CATCTGTCTGCAACCGCTCACTCAATGTAATGGAATGGAATGGAATGGAATGGAATGGA 1740  
Qy 1741 CCCCCAAAAGAAACCGGCTACACAGAGTGTAGTAAAGTCTTTTAAAGAAAGAACTGGAAGTTAA 1800  
Db 1741 CCCCCAAAAGAAACCGGCTACACAGAGTGTAGTAAAGTCTTTTAAAGAAAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTGCAAGTATGATGATAAAGCCAAAGGCTCTCTTCCCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTGCAAGTATGATGATAAAGCCAAAGGCTCTCTTCCCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCACTGTGACACAGGAAAGAAAGAA 1920  
Db 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCACTGTGACACAGGAAAGAAAGAA 1920  
Qy 1921 ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACTTACTAGAAAAAGTTCCCGGAGA 1980  
Db 1921 ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACTTACTAGAAAAAGTTCCCGGAGA 1980  
Qy 1981 CAAAGGATGCAAGAGAGGACAGATCAATATTTCCATCATGACCCCGGAGCTCACGTGT 2040  
Db 1981 CAAAGGATGCAAGAGAGGACAGATCAATATTTCCATCATGACCCCGGAGCTCACGTGT 2040  
Qy 2041 GCGCGCACACAGGGGCCCAAACTGAGGCTCCGAACTGCACTGCGGCTTGGTATTCCTCA 2100  
Db 2041 GCGCGCACACAGGGGCCCAAACTGAGGCTCCGAACTGCACTGCGGCTTGGTATTCCTCA 2100  
Qy 2101 AGGAAGCTGCAAGATTCGATGTGCCAACAGAGACAGGACCTGGAGGAAAGGCAAGGTGC 2160  
Db 2101 AGGAAGCTGCAAGATTCGATGTGCCAACAGAGACAGGACCTGGAGGAAAGGCAAGGTGC 2160  
Qy 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGCAGGATGCTCATCTTTCCGCTCA 2220  
Db 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGCAGGATGCTCATCTTTCCGCTCA 2220  
Qy 2221 TATTTCATGCTGGATGTGCGCATCCGAACTGACACCAAGAGAGAGGAGGCTTCCAG 2280  
Db 2221 TATTTCATGCTGGATGTGCGCATCCGAACTGACACCAAGAGAGAGGAGGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGAACTCTGGAGAGA 2324

## RESULT 4

US-09-903-216-3  
; Sequence 3, Application US/09903216  
; Patent No. US2002011481A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV2

```
; CURRENT APPLICATION NUMBER: US/09/903,216
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2324
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-903-216-3

Query Match      100.0%; Score 2324; DB 9; Length 2324;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCGTGCAATGGCCCGCAGCGTAAGAAATCCAAAGAGCAGCGGCAACAGCAGCAGCAGCG 60
DB 1 CGGACCGTGCAATGGCCCGCAGCGTAAGAAATCCAAAGAGCAGCGGCAACAGCAGCAGCAGCG 60

QY 61 GCTCCGCGAGCGGTAGCAGCGTGGCGCAGCAGCGCCCGGGGCCCGGAGAGAGACAA 120
DB 61 GCTCCGCGAGCGGTAGCAGCGTGGCGCAGCAGCGCCCGGGGCCCGGAGAGAGACAA 120

QY 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGACTCTCGGGAACTTTCATTCTTCAGT 180
DB 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGACTCTCGGGAACTTTCATTCTTCAGT 180

QY 181 GGTTCATGGTGAATGTCATGTCGTGGCGCTGGACATCTGTAGCTGTCGTTTGGTTGATC 240
DB 181 GGTTCATGGTGAATGTCATGTCGTGGCGCTGGACATCTGTAGCTGTCGTTTGGTTGATC 240

QY 241 TTGTTGACTATGAGGAAGTCTAGGAAATAGGAATCTATGATCTGATGCTGATGAG 300
DB 241 TTGTTGACTATGAGGAAGTCTAGGAAATAGGAATCTATGATCTGATGCTGATGAG 300

QY 301 ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTTCAGAGC 360
DB 301 ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTTCAGAGC 360

QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACACTGAGCCCGAGAGCAGGTTCTGTGG 420
DB 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGCAGGTTCTGTGG 420

QY 421 AGSCAGAACCCAGAAATTCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480
DB 421 AGSCAGAACCCAGAAATTCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480

QY 481 AAATGGTACACGCAAGCATGTTGAGGAGAGAGACTTGCAACAGAGATGGACCCACAG 540
DB 481 AAATGGTACACGCAAGCATGTTGAGGAGAGAGACTTGCAACAGAGATGGACCCACAG 540

QY 541 GAGAACCAACAGAGGATGATGATTTCTTATGGCGACTGATCTAGATGATGATTTG 600
DB 541 GAGAACCAACAGAGGATGATGATTTCTTATGGCGACTGATCTAGATGATGATTTG 600

QY 601 AGACCTCGAACTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAAGAGA 660
DB 601 AGACCTCGAACTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAAGAGA 660

QY 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTCGAGCAGGAAATCCAG 720
DB 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTCGAGCAGGAAATCCAG 720

QY 721 ATTCCAGTGAACAGTAGTAGAGATGAAGATTGCACCATCATACAGATGATGTAACAT 780
DB 721 ATTCCAGTGAACAGTAGTAGAGATGAAGATTGCACCATCATACAGATGATGTAACAT 780

QY 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840
DB 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840

QY 841 CAGAAGTAAGTGTCCCTCCCTGAGGATAATCTCTGTAGAGATTCACAGGTAATTTGAGAG 900
DB 841 CAGAAGTAAGTGTCCCTCCCTGAGGATAATCTCTGTAGAGATTCACAGGTAATTTGAGAG 900

841 CAGAAGTAAGTGTCCCTCCCTGAGGATAATCTCTGTAGAGATTCACAGGTAATTTGAGAG 900
901 AAGTAAGCAATTTTCTGTGGAGAAACAGCAGCAAGTACCAACAGAAAAACAATAGAAAAA 960
901 AAGTAAGCAATTTTCTGTGGAGAAACAGCAGCAAGTACCAACAGAAAAACAATAGAAAAA 960
961 CAGATGATCCAGAACAAAAAGCTTTAAGAAAAAGAGCCCTAAACTTTTAAAAAT 1020
961 CAGATGATCCAGAACAAAAAGCTTTAAGAAAAAGAGCCCTAAACTTTTAAAAAT 1020
1021 TTGATAAGACTATTAAGCTGAATGCTGCTCAGAAAAAATCCGTAAAGGAAAAA 1080
1021 TTGATAAGACTATTAAGCTGAATGCTGCTCAGAAAAAATCCGTAAAGGAAAAA 1080
1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCCCTCAGAGTCCAGAG 1140
1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCCCTCAGAGTCCAGAG 1140
1141 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGAGTAATGAGG 1200
1141 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGAGTAATGAGG 1200
1201 TGCTACGTGAGCCATCGACACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCAG 1260
1201 TGCTACGTGAGCCATCGACACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCAG 1260
1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320
1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320
1321 GAGGTTCCCTGCTTACCTCTGAGAGATTTAGTTCAACTATTTTCCCAATGATATCTTCTTAA 1380
1321 GAGGTTCCCTGCTTACCTCTGAGAGATTTAGTTCAACTATTTTCCCAATGATATCTTCTTAA 1380
1381 AAAATGACCTTGGCTGGGATACCTCTGTAGAGAGATATGCAATGCAAGAAAGTTT 1440
1381 AAAATGACCTTGGCTGGGATACCTCTGTAGAGAGATATGCAATGCAAGAAAGTTT 1440
1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGTCTTAAAGTCCATATGAGTTCA 1500
1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGTCTTAAAGTCCATATGAGTTCA 1500
1501 TCCTGAAGGACAGAACAAAAATTCCTGAGAGCATCCCATATTTAAAGAGAGAGATGAGAT 1560
1501 TCCTGAAGGACAGAACAAAAATTCCTGAGAGCATCCCATATTTAAAGAGAGAGATGAGAT 1560
1561 CCGGAGATCCCTGGCAGTGTGATGATGAGGATTTTATTTCCACCTGGGGGATGCCATGAGA 1620
1561 CCGGAGATCCCTGGCAGTGTGATGATGAGGATTTTATTTCCACCTGGGGGATGCCATGAGA 1620
1621 GGGTTGGGAAACAAAGAGGCAATATAAGTGTGATGAGTGGGCAACAAGAGAGACATTTTG 1680
1621 GGGTTGGGAAACAAAGAGGCAATATAAGTGTGATGAGTGGGCAACAAGAGAGACATTTTG 1680
1681 CATCTGTGCAACGCTCACTCTACATGTGATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 1740
1681 CATCTGTGCAACGCTCACTCTACATGTGATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 1740
1741 CCCCAAGAAACCGGGCTTACAGAGTGTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAA 1800
1741 CCCCAAGAAACCGGGCTTACAGAGTGTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAA 1800
1801 TCCGAGATGAAGGCTTGCAGTGTGATGATGAAGCAAGAGTCTCTTCTGCTGAGGATG 1860
1801 TCCGAGATGAAGGCTTGCAGTGTGATGATGAAGCAAGAGTCTCTTCTGCTGAGGATG 1860
1861 AAAACCTGAGGAAAGAGGAGTGGAGGAGTTCAGCTGTGGCAGCAAGAGAGAGAGAGAG 1920
1861 AAAACCTGAGGAAAGAGGAGTGGAGGAGTTCAGCTGTGGCAGCAAGAGAGAGAGAGAG 1920
1921 ATGAAAATGCTCAGAAAGAGGAGTCTCTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980
```

Db 1921 ATGAATAATGCTGCAAGAGGAGCTCTTAATAAAGCTGTACCTTACTAGTAAAGAGTTCCTCCGAGA 1980  
Qy 1981 CAACAGGATGAGAGAGAGAGATCAAAATATTCATATGACACCCCGGAGCTCAGCTGT 2040  
Db 1981 CAACAGGATGAGAGAGAGAGATCAAAATATTCATATGACACCCCGGAGCTCAGCTGT 2040  
Qy 2041 GCGCGCACACAGAGGCGCCACAACTGAGGCTCCGAATGACCTGGGCTTGGTGAATCCCA 2100  
Db 2041 GCGCGCACACAGAGGCGCCACAACTGAGGCTCCGAATGACCTGGGCTTGGTGAATCCCA 2100  
Qy 2101 AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACAGGACCTGGGAGGAGGCAAGGTGC 2160  
Db 2101 AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACAGGACCTGGGAGGAGGCAAGGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTGAGCAGGAGGTATGCGAGGATGCTCATCTTCCGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTGAGCAGGAGGTATGCGAGGATGCTCATCTTCCGCTGA 2220  
Qy 2221 TATTATCGTGGATGTGTGCGATTCGGAATGACACACAGAGAGAGCGAGCTTCCAG 2280  
Db 2221 TATTATCGTGGATGTGTGCGATTCGGAATGACACACAGAGAGAGCGAGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTCGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTCGAGAGA 2324

## RESULT 5

US-09-903-199-3  
; Sequence 3, Application US/09903199  
; Patent No. US20020122802A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV4  
; CURRENT APPLICATION NUMBER: US/09/903,199  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-199-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCAATGCGCCAGCGTAAGAATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCAATGCGCCAGCGTAAGAATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGGCAGCGGTAGCAGCAGGTGCGGCGAGCAGCAGCCCGGGGCGGAGAGAGACAA 120  
Db 61 GCTCCGGCAGCGGTAGCAGCAGGTGCGGCGAGCAGCAGCCCGGGGCGGAGAGAGACAA 120  
Qy 121 AGCATGGAGGACACAAGATGGAGGAAAGCGGACTCTCGGGAACCTTCATTCTTCACGT 180  
Db 121 AGCATGGAGGACACAAGATGGAGGAAAGCGGACTCTCGGGAACCTTCATTCTTCACGT 180  
Qy 181 GGTATTATGGTATTCATTCGCGGCTCGACATCTGTAGCTGTGCTTTGGTTTGTATC 240  
Db 181 GGTATTATGGTATTCATTCGCGGCTCGACATCTGTAGCTGTGCTTTGGTTTGTATC 240  
Qy 241 TTGTTGATATGAGAGTTCTAGGAAACTAGGAATCTATGATGCTGTGATGGGAG 300  
Db 241 TTGTTGATATGAGAGTTCTAGGAAACTAGGAATCTATGATGCTGTGATGGGAG 300

Qy 301 ATTTTGATGTGGATGATGCGCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
Db 301 ATTTTGATGTGGATGATGCGCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTCTGTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTCTGTGG 420  
Qy 421 AGGAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGAGACACATGTTGAGGGAGAGAGCTTGCACACAGAGATGGACCCACAG 540  
Db 481 AAATGGTACACGAGACACATGTTGAGGGAGAGAGCTTGCACACAGAGATGGACCCACAG 540  
Qy 541 GAGAAACCAACAGAGAGATGATGATTTCTTATGGCGACTGATGTAGATGATAGATTTG 600  
Db 541 GAGAAACCAACAGAGAGATGATGATTTCTTATGGCGACTGATGTAGATGATAGATTTG 600  
Qy 601 AGACCTCGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCCAGTGGAAAGA 660  
Db 601 AGACCTCGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCCAGTGGAAAGA 660  
Qy 661 CAGTTTCAAGAGCTGTAAATCAGGATATGGAAGAGATGATGCTGTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGAGCTGTAAATCAGGATATGGAAGAGATGATGCTGTGAGCAGGAAATCCAG 720  
Qy 721 ATTCAGTGAACCAAGTAGTAGAGATGAAGATTTGCCACATGATACAGATGATGAAT 780  
Db 721 ATTCAGTGAACCAAGTAGTAGAGATGAAGATTTGCCACATGATACAGATGATGAAT 780  
Qy 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAATCA 840  
Qy 841 CAGAAGTAATCTGCTCCCGCTGAGGATTAATCTGTAGAAAGATTCACAGGTAAATGTAGAAG 900  
Db 841 CAGAAGTAATCTGCTCCCGCTGAGGATTAATCTGTAGAAAGATTCACAGGTAAATGTAGAAG 900  
Qy 901 AAGTAAGCATTTTCTGTGGAGAACAGCAGGAGTACACACAGAAACAAATAGAAAAA 960  
Db 901 AAGTAAGCATTTTCTGTGGAGAACAGCAGGAGTACACACAGGAGTACACACAGAAACAAATAGAAAAA 960  
Qy 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGACCTTAAACTTTTAAATAAT 1020  
Db 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGACCTTAAACTTTTAAATAAT 1020  
Qy 1021 TTGATAAGACTATTAAAGCTGAACCTTGATGTGCGAGAAAAAACTCCGTAAAGGGGAAAAA 1080  
Db 1021 TTGATAAGACTATTAAAGCTGAACCTTGATGTGCGAGAAAAAACTCCGTAAAGGGGAAAAA 1080  
Qy 1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTTCAGAGTCCACGAG 1140  
Qy 1141 CAAGATATGGAGAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
Db 1141 CAAGATATGGAGAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
Qy 1201 TGCTACGTGGAGCATCGAGACCTACCAAGAGGTGGCCAGCTTACCTGATGTCCCTGCGAG 1260  
Db 1201 TGCTACGTGGAGCATCGAGACCTACCAAGAGGTGGCCAGCTTACCTGATGTCCCTGCGAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAAATTTCTAGGTCAATAGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAAATTTCTAGGTCAATAGA 1320  
Qy 1321 GAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTCAACTATTTCCCAATGATCTTCTTAA 1380  
Db 1321 GAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTCAACTATTTCCCAATGATCTTCTTAA 1380

Qy	1381	AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATAATGACAATGCAAGAAAGTTT	1441
Db	1381	AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATAATGACAATGCAAGAAAGTTT	1440
Qy	1441	ATGAAGAGGTGCTCAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA	1500
Db	1441	ATGAAGAGGTGCTCAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA	1500
Qy	1501	TCCTGAAAGGCACAGAAACAAAATTCCTGAGAGCATCCCATATTTAAAGAAAGGAATAGAAAT	1560
Db	1501	TCCTGAAGGCACAGAAACAAAATTCCTGAGAGCATCCCATATTTAAAGAAAGGAATAGAAAT	1560
Qy	1561	CCGAGATCTCTGGCCTGATGATGGGAGATTTTATTTTCACCTGGGGGATGCCCATGCGAGA	1620
Db	1561	CCGAGATCTCTGGCCTGATGATGGGAGATTTTATTTTCACCTGGGGGATGCCCATGCGAGA	1620
Qy	1621	GGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACACCTTTG	1680
Db	1621	GGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACACCTTTG	1680
Qy	1681	CATCTGTCTGGCAACCGCTCACTCTACAATGTGAATGGACTGAAAGCACAGCCTTGGTGGGA	1740
Db	1681	CATCTGTCTGGCAACCGCTCACTCTACAATGTGAATGGACTGAAAGCACAGCCTTGGTGGGA	1740
Qy	1741	CCCCAAAGAAACCGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA	1800
Db	1741	CCCCAAAGAAACCGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA	1800
Qy	1801	TCCGAGATGAAGGCTTGCAGTGTATGATTAAGCCAAAGTCTCTTCTGCCTGAGAGTG	1860
Db	1801	TCCGAGATGAAGGCTTGCAGTGTATGATTAAGCCAAAGTCTCTTCTGCCTGAGAGTG	1860
Qy	1861	AAAACTTGAGGGAACAAAGGGGACCTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA	1920
Db	1861	AAAACTTGAGGGAACAAAGGGGACCTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA	1920
Qy	1921	ATGAAAATGCTGCAAAAGGAGCTCCTAAAACTGTACTATGAAAAAGTTTCCCCGAGA	1980
Db	1921	ATGAAAATGCTGCAAAAGGAGCTCCTAAAACTGTACTATGAAAAAGTTTCCCCGAGA	1980
Qy	1981	CAACAGATGCAGAAAGGACAGATCAATATTTCCATCATGCACCCCGGAGCTCACCTGT	2040
Db	1981	CAACAGATGCAGAAAGGACAGATCAATATTTCCATCATGCACCCCGGAGCTCACCTGT	2040
Qy	2041	GGCCGCACACAGGGGCCACAAACTGCAGGCTCCGAATGCACCTGGGCTTGGTGAATCCCA	2100
Db	2041	GGCCGCACACAGGGGCCACAAACTGCAGGCTCCGAATGCACCTGGGCTTGGTGAATCCCA	2100
Qy	2101	AGGAAGGCTGCAAGATTGATGTGCCAAACGAGACCCAGGACTGGGAGGAAGGCAAGTGC	2160
Db	2101	AGGAAGGCTGCAAGATTGATGTGCCAAACGAGACCCAGGACTGGGAGGAAGGCAAGTGC	2160
Qy	2161	TCATCTTTGATGACTCTCTTTGAGCACAGAGTATGGCAGGATGCGCTCATCTTTCCGGCTGA	2220
Db	2161	TCATCTTTGATGACTCTCTTTGAGCACAGAGTATGGCAGGATGCGCTCATCTTTCCGGCTGA	2220
Qy	2221	TATTTCATGTGATGTGTGGCATCCGGAACTGACACCAACAGCAGACGACGACCTTCCAG	2280
Db	2221	TATTTCATGTGATGTGTGGCATCCGGAACTGACACCAACAGCAGACGACGACCTTCCAG	2280
Qy	2281	CAATTTAGCATGAATTCATGCAAGCTTTGGGAACTCTGGAGAGA	2324
Db	2281	CAATTTAGCATGAATTCATGCAAGCTTTGGGAACTCTGGAGAGA	2324

## RESULT 6

```

US-09-903-023-3
; Sequence 3, Application US/09903023
; Patent No. US2002014621A1
; GENERAL INFORMATION:
; APPLICANT: wands, Jack R.
; APPLICANT: de la Monte, Suzanne M
; APPLICANT: Ince, Nedim

```

Db 781 ACCAAGTCTATGAGAAACAAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
 Qy 841 CAGAAATGCTCCCTCCCTGAGGATAATCTGTGAGAGATTCACAGGTAAATGTTGAGAAG 900  
 Db 841 CAGAAATGCTCCCTCCCTGAGGATAATCTGTGAGAGATTCACAGGTAAATGTTGAGAAG 900  
 Qy 901 AAGTAAGCATTTTCTGTGTGAAAGAACAGCAGCAGAGAGTACCAACGAGAAACAAATAGAAAAA 960  
 Db 901 AAGTAAGCATTTTCTGTGTGAAAGAACAGCAGCAGAGAGTACCAACGAGAAACAAATAGAAAAA 960  
 Qy 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAGAACCTTAAATCTTTTAAATTAAT 1020  
 Db 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAGAACCTTAAATCTTTTAAATTAAT 1020  
 Qy 1021 TTGATAAGCACTATTAAAGCTGAACCTTGTGCTGCAAGAAAACCTCGTAAAGGGGAAAAA 1080  
 Db 1021 TTGATAAGCACTATTAAAGCTGAACCTTGTGCTGCAAGAAAACCTCGTAAAGGGGAAAAA 1080  
 Qy 1081 TTGAGAAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
 Db 1081 TTGAGAAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
 Qy 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAAATGAGG 1200  
 Db 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAAATGAGG 1200  
 Qy 1201 TGCTACGTGGAGCCATCGACACCTACCAAGAGTGGCCAGCTACCTGATGTCCTCGAG 1260  
 Db 1201 TGCTACGTGGAGCCATCGACACCTACCAAGAGTGGCCAGCTACCTGATGTCCTCGAG 1260  
 Qy 1261 ACTGCTGAGCTGAGTTTGAAGCTGCTGAGCAGGCAACAAATTTCTAGGTCAATATGA 1320  
 Db 1261 ACTGCTGAGCTGAGTTTGAAGCTGCTGAGCAGGCAACAAATTTCTAGGTCAATATGA 1320  
 Qy 1321 GAGGTTCCCTCTTACCCTGACAGATATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
 Db 1321 GAGGTTCCCTCTTACCCTGACAGATATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
 Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
 Db 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
 Qy 1441 ATGAGAGTCTGAGTGTGACACCTAATGATGCTTGTCTAAAGTCCATATGCGTCA 1500  
 Db 1441 ATGAGAGTCTGAGTGTGACACCTAATGATGCTTGTCTAAAGTCCATATGCGTCA 1500  
 Qy 1501 TCCTGAAGGCAAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAAGGAAT 1560  
 Db 1501 TCCTGAAGGCAAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAAGGAAT 1560  
 Qy 1561 CCGAGATCTCTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATCAGA 1620  
 Db 1561 CCGAGATCTCTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATCAGA 1620  
 Qy 1621 GGGTTGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGGACACTTGG 1680  
 Db 1621 GGGTTGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGGACACTTGG 1680  
 Qy 1681 CATCTGCTGCAACGCTCACTTCAATGTGAATGGAATGGAATGGAATGGAATGGAATGGA 1740  
 Db 1681 CATCTGCTGCAACGCTCACTTCAATGTGAATGGAATGGAATGGAATGGAATGGAATGGA 1740  
 Qy 1741 CCCCAGAAAGAACGGGTACACAGATTTAGTAAGTCTTTTAAAGAAACCTGGAAGTTAA 1800  
 Db 1741 CCCCAGAAAGAACGGGTACACAGATTTAGTAAGTCTTTTAAAGAAACCTGGAAGTTAA 1800  
 Qy 1801 TCCGAGATGAAGGCTTTCAGATGATGATAAAGCAAAAGGCTCTTCTCCTCGTGGAGT 1860  
 Db 1801 TCCGAGATGAAGGCTTTCAGATGATGATAAAGCAAAAGGCTCTTCTCCTCGTGGAGT 1860  
 Qy 1861 AAAACCTGAGGAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAGGAGAGAA 1920  
 Db 1861 AAAACCTGAGGAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAGGAGAGAA 1920

RESULT 7

US-09-436-184-3  
 ; Sequence 3, Application US/09436184  
 ; Publication No. US20030031670A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wands, Jack R.  
 ; APPLICANT: de la Monte, Suzanne M.  
 ; APPLICANT: Ince, Nedim  
 ; APPLICANT: Carlson, Rolf I.  
 ; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
 ; FILE REFERENCE: R. I. Hoep. - Malignant Neoplasms  
 ; CURRENT APPLICATION NUMBER: US/09/436,184  
 ; CURRENT FILING DATE: 1999-11-08  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 2324  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-436-184-3

Query Match 100.0%; Score 2324; DB 10; Length 2324;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCATGGCCCGCCAGCGTAAGATGCCAAGCAGCAGCGCAACAGCAGCAGCAGCG 60  
 Db 1 CGGACCGTGCATGGCCCGCCAGCGTAAGATGCCAAGCAGCAGCGCAACAGCAGCAGCAGCG 60  
 Qy 61 GCTCCGGCAGCGGTAGCAGCAGGTGCGGGCAGCAGCAGCCCCCGGGCCCGGAGAGACAA 120  
 Db 61 GCTCCGGCAGCGGTAGCAGCAGGTGCGGGCAGCAGCAGCCCCCGGGCCCGGAGAGACAA 120  
 Qy 121 AGCATGGAGGACAAAGATGGAGGAAAGGCGGACTCTCGGGAACCTTCATTCTTCCAGT 180  
 Db 121 AGCATGGAGGACAAAGATGGAGGAAAGGCGGACTCTCGGGAACCTTCATTCTTCCAGT 180  
 Qy 181 GGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
 Db 181 GGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
 Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGCTGCTGCTGCTG 300

Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGATGGTATGGAG 300  
Qy 301 ATTTTGTATGTGATGATGCCAAAGTTTATTATAGAGCTTAAAGAGAGATCTACTTTCAGAGC 360  
Db 301 ATTTTGTATGTGATGATGCCAAAGTTTATTATAGAGCTTAAAGAGAGATCTACTTTCAGAGC 360  
Qy 361 CAGCAGTCCCAGGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGTGG 420  
Db 361 CAGCAGTCCCAGGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGTGG 420  
Qy 421 AGGCGAAGACCCAGAGATTCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGCGAAGACCCAGAGATTCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGAGAACTGTTGAGGAGAGAGACTTGCACAAAGAGATGGACCCACAG 540  
Db 481 AAATGGTACACGAGAACTGTTGAGGAGAGAGACTTGCACAAAGAGATGGACCCACAG 540  
Qy 541 GAGAACCAACAAAGAGATGATGAGTTTCTTATGGCGACTGATGTAGATGATAGATTG 600  
Db 541 GAGAACCAACAAAGAGATGATGAGTTTCTTATGGCGACTGATGTAGATGATAGATTG 600  
Qy 601 AGACCTCGAACCCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACAGTGAAGAGA 660  
Db 601 AGACCTCGAACCCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACAGTGAAGAGA 660  
Qy 661 CAGTTTCAAGACTGTAACTCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGACTGTAACTCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACCAAGTAGTAGAAGATGACCATGATACAGATGATGTAAACAT 780  
Db 721 ATTCCAGTGAACCAAGTAGTAGAAGATGACCATGATACAGATGATGTAAACAT 780  
Qy 781 ACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840  
Qy 841 CAGAGTAACTGCTCCCTCAGGATATCTGTAGAGATTCACAGTAAATGTAGAG 900  
Db 841 CAGAGTAACTGCTCCCTCAGGATATCTGTAGAGATTCACAGTAAATGTAGAG 900  
Qy 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGTACCACGAAACAAATAGAAAA 960  
Db 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGTACCACGAAACAAATAGAAAA 960  
Qy 961 CAGATGATCCAGAACAAAGCAAAAGTTAGAAAAGAACCTTAACCTTTAAATAAT 1020  
Db 961 CAGATGATCCAGAACAAAGCAAAAGTTAGAAAAGAACCTTAACCTTTAAATAAT 1020  
Qy 1021 TTGATAAGCACTATTAAAGCTGAACCTTGATGTCGAGAAAACCTCCGTAAAAAGGGAAAAA 1080  
Db 1021 TTGATAAGCACTATTAAAGCTGAACCTTGATGTCGAGAAAACCTCCGTAAAAAGGGAAAAA 1080  
Qy 1081 TTGAGAAAGCAGTGAATGCAITTAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
Db 1081 TTGAGAAAGCAGTGAATGCAITTAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
Qy 1141 CAAGATATGGGAAGCGCAGTGTGAGATGATTTGGCTGAGAGAGAGAGATTAATGAGG 1200  
Db 1141 CAAGATATGGGAAGCGCAGTGTGAGATGATTTGGCTGAGAGAGAGAGATTAATGAGG 1200  
Qy 1201 TGCTACGTGGAGCCATCCAGACCTACCAAGAGTGGCCAGCTACCTCATGTCCCTGCAG 1260  
Db 1201 TGCTACGTGGAGCCATCCAGACCTACCAAGAGTGGCCAGCTACCTCATGTCCCTGCAG 1260  
Qy 1261 ACCTGCTGAAGCTGATTTGAAGCGCTCGCTCAGACAGGCAACAAATTTCTAGTCAATATGA 1320  
Db 1261 ACCTGCTGAAGCTGATTTGAAGCGCTCGCTCAGACAGGCAACAAATTTCTAGTCAATATGA 1320  
Qy 1321 GAGGTTCCCTGCTTACCTGCGAGAGATTAATGTTCAACTATTTCCCAATGATACCTTCTTAA 1380  
Db 1321 GAGGTTCCCTGCTTACCTGCGAGAGATTAATGTTCAACTATTTCCCAATGATACCTTCTTAA 1380

Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATAATGACAAATGCAAGAAAGTTT 1440  
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATAATGACAAATGCAAGAAAGTTT 1440  
Qy 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATATATGGCTTCA 1500  
Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATATATGGCTTCA 1500  
Qy 1501 TCCTGAAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGATAGAAT 1560  
Db 1501 TCCTGAAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGATAGAAT 1560  
Qy 1561 CCGGAGATCTCTGGCAGCTGATGGGAGATTTTATTTTCCACCTGGGGGATGCCATGCGA 1620  
Db 1561 CCGGAGATCTCTGGCAGCTGATGGGAGATTTTATTTTCCACCTGGGGGATGCCATGCGA 1620  
Qy 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGGACACATTTG 1680  
Db 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGGACACATTTG 1680  
Qy 1681 CATCTGTCTGCAACGCTCACTCTACAAATGTGAATGACTGAAAGACAGCCTTGGTGA 1740  
Db 1681 CATCTGTCTGCAACGCTCACTCTACAAATGTGAATGACTGAAAGACAGCCTTGGTGA 1740  
Qy 1741 CCCCAGAAAGAAACGGGCTCACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
Db 1741 CCCCAGAAAGAAACGGGCTCACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTCAGTGTATGATTAAGCAAAAGTCTCTTCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTCAGTGTATGATTAAGCAAAAGTCTCTTCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAAGGGGACTGGAGCCAGTTTACGCTGTGCGACAGCAAGAAAGAA 1920  
Db 1861 AAAACCTGAGGAAAAGGGGACTGGAGCCAGTTTACGCTGTGCGACAGCAAGAAAGAA 1920  
Qy 1921 ATGAAAAATGCTGCAAAAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAGA 1980  
Db 1921 ATGAAAAATGCTGCAAAAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAGA 1980  
Qy 1981 CAACAGATGTCAGAAAGAGGACAGATCAAAATTTCCATCATGCAACCCCGGACTCACGTGT 2040  
Db 1981 CAACAGATGTCAGAAAGAGGACAGATCAAAATTTCCATCATGCAACCCCGGACTCACGTGT 2040  
Qy 2041 GGCCGCAACAGGCCCCACAAACTGCAAGCTCCGAAATGCACTGGGCTTGGTGAATCCCA 2100  
Db 2041 GGCCGCAACAGGCCCCACAAACTGCAAGCTCCGAAATGCACTGGGCTTGGTGAATCCCA 2100  
Qy 2101 AGGAAGGCTGCAAGATTCGATGTGCCAAACGAGACCTGGGAGGAAGGCAAGGTGC 2160  
Db 2101 AGGAAGGCTGCAAGATTCGATGTGCCAAACGAGACCTGGGAGGAAGGCAAGGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTGAGCAGCAGGATGATGGCAGGATGCCTCATCTTTCGGGTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTGAGCAGCAGGATGATGGCAGGATGCCTCATCTTTCGGGTGA 2220  
Qy 2221 TATTTCATGCTGGATGTGTCATCCGAACTGACCAACAGCAGACGACGAGCTTCCAG 2280  
Db 2221 TATTTCATGCTGGATGTGTCATCCGAACTGACCAACAGCAGACGACGAGCTTCCAG 2280  
Qy 2281 CAATTTAGCTGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCTGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324

RESULT 8  
US-10-084-817-63  
; Sequence 63, Application US/10084817  
; Publication No. US20030119009A1  
; GENERAL INFORMATION:  
; APPLICANT: Susan Stuart  
; APPLICANT: Jed G. Nuchtern



APPLICANT: Sharon E. Plon  
APPLICANT: Jason M. Shonhet  
TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
FILE REFERENCE: PA-0046 US  
CURRENT APPLICATION NUMBER: US/10/084,817  
CURRENT FILING DATE: 2002-02-25  
PRIOR APPLICATION NUMBER: 60/270,784  
PRIOR FILING DATE: 2001-02-23  
NUMBER OF SEQ ID NOS: 365  
SOFTWARE: PERL Program  
SEQ ID NO 63  
LENGTH: 5358  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. US20030119009A1 089172.13  
US-10-084-817-63

Query Match 93.4%; Score 2171; DB 15; Length 5358;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	CGGACCGTGCATGGCCAGCGTAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60
DB	258	CGGACCGTGCATGGCCAGCGTAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 317
QY	61	GCTCCGGCAGCGGTAGCAGCAGTCCGGCAGCAGCAGCCCGGGCCCGGAGAGACAA 120
DB	318	GCTCCGGCAGCGGTAGCAGCAGTCCGGCAGCAGCAGCCCGGGCCCGGAGAGACAA 377
QY	121	AGCATGGAGGACACAGAATGGAGGAAAGCGGACTCTCGGGAACTTCATCTTCACGT 180
DB	378	AGCATGGAGGACACAGAATGGAGGAAAGCGGACTCTCGGGAACTTCATCTTCACGT 437
QY	181	GCTTATGGTGAATGCTGCGGGCTCTGGACACTCTGTAGCTGTGCTGTTGGTTGATC 240
DB	438	GCTTATGGTGAATGCTGCGGGCTCTGGACACTCTGTAGCTGTGCTGTTGGTTGATC 497
QY	241	TGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300
DB	498	TGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 557
QY	301	ATTGTGATGGATGATCCAAAGTTTATTAGGACTTAAAGAGATCTACTTCAGAGC 360
DB	558	ATTGTGATGGATGATCCAAAGTTTATTAGGACTTAAAGAGATCTACTTCAGAGC 617
QY	361	CAGCATCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG 420
DB	618	CAGCATCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG 677
QY	421	AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480
DB	678	AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 737
QY	481	AAATGGTACACGACAGACATGTTGAGGAGAGAGCTTGCACAGAGAGATGACCCACAG 540
DB	738	AAATGGTACACGACAGACATGTTGAGGAGAGAGCTTGCACAGAGAGATGACCCACAG 797
QY	541	GAGAACCAACAGAGGATGATGAGTTCTTATGGCACTGATGTAGATGATGATTTG 600
DB	798	GAGAACCAACAGAGGATGATGAGTTCTTATGGCACTGATGTAGATGATGATTTG 857
QY	601	AGACCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTTTACCACTGGGAAGA 660
DB	858	AGACCTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTTTACCACTGGGAAGA 917
QY	661	CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720
DB	918	CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 977
QY	721	ATTCCAGTGAACAGTAGTAGAGATGAAGATTCACCATGATACAGATGATGATCAAT 780

DB	978	ATTCAGTGAACAGTAGTAGAGATGAAGAGTTGCACCATGATACAGATGATGTAACAT 1037
QY	781	ACCAAGTCTATGAGGAAACAAGAGTATATGAACCTCTTAGAAATGAAGGATGAGAATCA 840
DB	1038	ACCAAGTCTATGAGGAAACAAGAGTATATGAACCTCTTAGAAATGAAGGATGAGAATCA 1097
QY	841	CAGAAGTAACTGCTCCCTCGAGGATAATCTCTGTAGAAGATTCAAGGATTAATTTGTAAG 900
DB	1098	CAGAAGTAACTGCTCCCTCGAGGATAATCTCTGTAGAAGATTCAAGGATTAATTTGTAAG 1157
QY	901	AACTAAGCATTTTTCTGTGGAAGAACAGCAGGAAGTACCACAGAAACAAATAGAAAA 960
DB	1158	AACTAAGCATTTTTCTGTGGAAGAACAGCAGGAAGTACCACAGAAACAAATAGAAAA 1217
QY	961	CAGATGATCCAGACAAAGAAAGTAAAGAAAGAGCGCTTAAACTTTTAATAAT 1020
DB	1218	CAGATGATCCAGACAAAGAAAGTAAAGAAAGAGCGCTTAAACTTTTAATAAT 1277
QY	1021	TTGATAAGACTATTAAAGCTGAATCTGTGCGAGAAAACTCCGTAAAGGGGAAAA 1080
DB	1278	TTGATAAGACTATTAAAGCTGAATCTGTGCGAGAAAACTCCGTAAAGGGGAAAA 1337
QY	1081	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAG 1140
DB	1338	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAG 1397
QY	1141	CAAGATATGGGAAGGCGCAGTGTAGGATGATTTGCGCTGAGAAGAGGAAGTAATGAGG 1200
DB	1398	CAAGATATGGGAAGGCGCAGTGTAGGATGATTTGCGCTGAGAAGAGGAAGTAATGAGG 1457
QY	1201	TGCTAGTGGAGCCATCGAGACCTACAAAGAGGTGGCCAGCTACTGTGTCCTGCGAG 1260
DB	1458	TGCTAGTGGAGCCATCGAGACCTACAAAGAGGTGGCCAGCTACTGTGTCCTGCGAG 1517
QY	1261	ACCTGCTGAAGCTGAGTTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGGTCTATGA 1320
DB	1518	ACCTGCTGAAGCTGAGTTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGGTCTATGA 1577
QY	1321	GAGGTTCCCTGCTTACCCTCGAGATTAGTTTCAACTATTTCCCAATGATCTTCTCTAA 1380
DB	1578	GAGGTTCCCTGCTTACCCTCGAGATTAGTTTCAACTATTTCCCAATGATCTTCTCTAA 1637
QY	1381	AAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGAACAATGAAGAAATTT 1440
DB	1638	AAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGAACAATGAAGAAATTT 1697
QY	1441	ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATATGCGCTCA 1500
DB	1698	ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATATGCGCTCA 1757
QY	1501	TCCTGAAGGCAAGAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAAT 1560
DB	1758	TCCTGAAGGCAAGAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAAT 1817
QY	1561	COGAGATCCCTGGCAGCTGATGAGGAGATTTTATTTCCACTCGGCGATGCCATGCGAGA 1620
DB	1818	COGAGATCCCTGGCAGCTGATGAGGAGATTTTATTTCCACTCGGCGGATGCCATGCGAGA 1877
QY	1621	GGTTTGGGAACAAGAGGCATATAAGTGTATGAGCTTGGGCAACAAGAGGACACTTTG 1680
DB	1878	GGTTTGGGAACAAGAGGCATATAAGTGTATGAGCTTGGGCAACAAGAGGACACTTTG 1937
QY	1681	CATCTGTCTGGCAACCGCTCACTCTCAATGTGAATGGAATGGAATGGAATGGAATGGA 1740
DB	1938	CATCTGTCTGGCAACCGCTCACTCTCAATGTGAATGGAATGGAATGGAATGGAATGGA 1997
QY	1741	CCCCAAAAGAAAACGGGCTACACAGAGTGTAGTAAAGTCTTTAGAAAGAAAACCTGGAAGTTAA 1800
DB	1998	CCCCAAAAGAAAACGGGCTACACAGAGTGTAGTAAAGTCTTTAGAAAGAAAACCTGGAAGTTAA 2057
QY	1801	TCCGAGATGAGGCTTGGAGTGAATGAAGCCAAAGGCTCTTCTCCCTCGGAGGATG 1860
DB	2058	TCCGAGATGAGGCTTGGAGTGAATGAAGCCAAAGGCTCTTCTCCCTCGGAGGATG 2117



Qy 1861 AAAACCTGAGGGAAGGAGGAGCTGGAGCCAGTTACGCTGTGGCAGCAAGGAAGAA 1920  
Db 2118 AAAACCTGAGGGAAGGAGGAGCTGGAGCCAGTTACGCTGTGGCAGCAAGGAAGAA 2177  
Qy 1921 ATGAATATGCTGCAAGGAGCTCTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
Db 2178 ATGAATATGCTGCAAGGAGCTCTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 2237  
Qy 1981 CAACAGGATGAGAGGAGCAGATCAATATTCATCATGACACCCGGGACTCACGTGT 2040  
Db 2238 CAACAGGATGAGAGGAGCAGATCAATATTCATCATGACACCCGGGACTCACGTGT 2297  
Qy 2041 GGCCGCACACAGGCGCCCAAACTGCAAGCTCCGAATGCACTGGCTTGGTATCCCA 2100  
Db 2298 GGCCGCACACAGGCGCCCAAACTGCAAGCTCCGAATGCACTGGCTTGGTATCCCA 2357  
Qy 2101 AGGAAGCTGCAAGATTCGATGTGCAACGAGCAGCAGGACCTGGGAGGAGCAAGTGC 2160  
Db 2358 AGGAAGCTGCAAGATTCGATGTGCAACGAGCAGCAGGACCTGGGAGGAGCAAGTGC 2417  
Qy 2161 TCATCTTTGATGACTCCCTTTGAGCAGCAGGATGTCAGGATGCCTCATCTTTCCGGCTGA 2220  
Db 2418 TCATCTTTGATGACTCCCTTTGAGCAGCAGGATGTCAGGATGCCTCATCTTTCCGGCTGA 2477  
Qy 2221 TATTCATCGTGGATGTGTGGCATCCGGAATGACACCAAGCAGAGCGCAGCTTCCAG 2280  
Db 2478 TATTCATCGTGGATGTGTGGCATCCGGAATGACACCAAGCAGAGCGCAGCTTCCAG 2537  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
Db 2538 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2581

RESULT 9  
US-10-096-534-37  
; Sequence 37, Application US/10096534  
; Publication No. US20030166887A1  
; GENERAL INFORMATION:  
; APPLICANT: The Brigham and Women's Hospital, Inc.  
; APPLICANT: Yates, Karen  
; APPLICANT: Mizuno, Shuichi  
; APPLICANT: Glowacki, Julie  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF SKELETAL DEGENERATION CONDITIONS  
; FILE REFERENCE: B0801/7244/KA/ERP  
; CURRENT APPLICATION NUMBER: US/10/096.534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US 60/274,980  
; PRIOR FILING DATE: 2001-03-12  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 37  
; LENGTH: 2680  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-096-534-37

Query Match 38.5%; Score 895; DB 16; Length 2680;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 945; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCATATGCGCCAGCGTAAGATGCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCATATGCGCCAGCGTAAGATGCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGCGAGCGGTAGCAGGTGGCGCAGCAGCAGCGCCGCGCGGAGAGACAA 120  
Db 61 GCTCCGCGAGCGGTAGCAGGTGGCGCAGCAGCAGCGCCGCGCGGAGAGACAA 120  
Qy 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGAGCTCTCGGGAATCTTCTTCAGT 180  
Db 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGAGCTCTCAGGAATCTTCTTCAGT 180

Qy 181 GGTTTATGGTGATTCATTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
Db 181 GGTTTATGGTGATTCATTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGAAAACCTAGGAATCTATGATGCTGATGGTATGAG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGAAAACCTAGGAATCTATGATGCTGATGGTATGAG 300  
Qy 301 ATTTTGTATGGTATGCTCCAAAGTTTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC 360  
Db 301 ATTTTGTATGGTATGCTCCAAAGTTTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTCTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTCTGG 420  
Qy 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACTTCAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACTTCAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGCGAAGACATGTTGAGGGAAGAATTGCAACAAGAGATGGACCCACAG 540  
Db 481 AAATGGTACACGCGAAGACATGTTGAGGGAAGAATTGCAACAAGAGATGGACCCACAG 540  
Qy 541 GAGAACCAACAAGAGAGATGATGATTTCTTATGGCGACTGTATGATGATGATGATTTG 600  
Db 541 GAGAACCAACAAGAGAGATGATGATTTCTTATGGCGACTGTATGATGATGATGATTTG 600  
Qy 601 AGACCTTGGAACTGAAATGATCTCATGAAGAAACCGAGCATAGTTACACGCTGGAAGAGA 660  
Db 601 AGACCTTGGAACTGAAATGATCTCATGAAGAAACCGAGCATAGTTACACGCTGGAAGAGA 660  
Qy 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACAGTAGTAGAAGATGAAGATGTCACCATGATACAGATGATGTAACAT 780  
Db 721 ATTCCAGTGAACAGTAGTAGAAGATGAAGATGTCACCATGATACAGATGATGTAACAT 780  
Qy 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840  
Qy 841 CAGAAGTAAGTCTCCCTGAGGATATCTCTAGAGATTCAGGTAATTTGTAGAG 900  
Db 841 CAGAAGTAAGTCTCCCTGAGGATATCTCTAGAGATTCAGGTAATTTGTAGAG 900  
Qy 901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGGAGTACCACACAGA 946  
Db 901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGGAGTACCACACAGA 946

RESULT 10  
US-10-269-909-3  
; Sequence 3, Application US/10269909  
; Publication No. US20030180747A1  
; GENERAL INFORMATION:  
; APPLICANT: HRUBAN, RALPH H.  
; APPLICANT: ARGANI, PEDRAM  
; APPLICANT: IACOBIZIO-DONAHUE, CHRISTINE  
; APPLICANT: MAITRA, ANIRBAN  
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES  
; CURRENT APPLICATION NUMBER: US/10/269,909  
; CURRENT FILING DATE: 2003-10-11  
; PRIOR APPLICATION NUMBER: 60/328,609  
; PRIOR FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 60/332,754  
; PRIOR FILING DATE: 2001-11-19  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3





Db 406 CAAGAAGATGGACCCACAGGAGAACACAAAGAGGATGATGAGTTCTTATGGCGACT 465  
Qy 582 GATGTAGATGATAGATTTGAGACCTCGAACCTGGAATATCTCATGAGAAACCGAGCAT 641  
Db 466 GATGTAGATGATAGATTTGAGACCTCGAACCTGGAATATCTCATGAGAAACCGAGCAT 525  
Qy 642 AGTTACCAAGTGGAGAGACAGTTTCAAGACTGTATCAGGATATGGAAGATGATG 701  
Db 526 AGTTACCAAGTGGAGAGACAGTTTCAAGACTGTATCAGGATATGGAAGATGATG 585  
Qy 702 TCTGAGCAGGAAATCCAGATTCAGTGAACCAAGTATGAGATGAAAGATTGCACCAT 761  
Db 586 TCTGAGCAGGAAATCCAGATTCAGTGAACCAAGTATGAGATGAAAGATTGCACCAT 645  
Qy 762 GATACAGATGATGTAACATACCAAGTCTATGAGAACCAAGCAGTATATGAACCTCTAGAA 821  
Db 646 GATACAGATGATGTAACATACCAAGTCTATGAGAACCAAGCAGTATATGAACCTCTAGAA 705  
Qy 822 AATGAAGGATAGAAATCCAGAGTAAGT 852  
Db 706 AATGAAGGATAGAAATCCAGAGTAAGT 736

## RESULT 14

US-10-017-754-1791  
; Sequence 1791, Application US/10017754  
; Publication No. US20030054363A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Warnerakis, Margarita  
; APPLICANT: Carter, Derrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C18  
; CURRENT APPLICATION NUMBER: US/10/017,754  
; CURRENT FILING DATE: 2001-10-29  
; NUMBER OF SEQ ID NOS: 2004  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-754-1791

Query Match 27.5%; Score 640; DB 14; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 162 GGAATCTCATCTTTCACGTGGTTTATGGTATTCATTCGTCGGCGCTCGACATCTGTA 221  
Db 46 GGAATCTCATCTTTCACGTGGTTTATGGTATTCATTCGTCGGCGCTCGACATCTGTA 105  
Qy 222 GCTGTCGTTTGGTTTGTATCTGTTGACTATGAGGAAGTCTTAGGAAACTTAGGAATCTAT 281  
Db 106 GCTGTCGTTTGGTTTGTATCTGTTGACTATGAGGAAGTCTTAGGAAACTTAGGAATCTAT 165  
Qy 282 GATGCTGATGATGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 166 GATGCTGATGATGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 225  
Qy 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401  
Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 285  
Qy 402 GAGGAGCAGGTTCTCTGTGGAGGCAGAACCCCAAGATATCGAAGATGAGCAAGAACAA 461

Db 286 GAGGAGCAGGTTCTCTGTGGAGCAGAACCCCAAGATATCGAAGATGAGCAAGAACAA 345  
Qy 462 ATTCACTCCCTTCTCCATGAAATGGTACACGAGAACATGTTGAGGAGAGAGACTTGCAG 521  
Db 346 ATTCACTCCCTTCTCCATGAAATGGTACACGAGAACATGTTGAGGAGAGAGACTTGCAG 405  
Qy 522 CAAGAAGATGACACCCACAGAGAACCAACCAAGAGGATGATGATGATGATGATGATGATG 581  
Db 406 CAAGAAGATGACACCCACAGAGAACCAACCAAGAGGATGATGATGATGATGATGATGATG 465  
Qy 582 GATGTAGATGATGATTTGAGACCTCGAACCTGGAATATCTCATGAGAAACCGAGCAT 641  
Db 466 GATGTAGATGATGATTTGAGACCTCGAACCTGGAATATCTCATGAGAAACCGAGCAT 525  
Qy 642 AGTTACCAAGTGGAGAGACAGTTTCAAGACTGTATCAGGATATGGAAGATGATG 701  
Db 526 AGTTACCAAGTGGAGAGACAGTTTCAAGACTGTATCAGGATATGGAAGATGATG 585  
Qy 702 TCTGAGCAGGAAATCCAGATTCAGTGAACCAAGTATGAGATGAAAGATTGCACCAT 761  
Db 586 TCTGAGCAGGAAATCCAGATTCAGTGAACCAAGTATGAGATGAAAGATTGCACCAT 645  
Qy 762 GATACAGATGATGTAACATACCAAGTCTATGAGAACCAAGCAGTATATGAACCTCTAGAA 821  
Db 646 GATACAGATGATGTAACATACCAAGTCTATGAGAACCAAGCAGTATATGAACCTCTAGAA 705  
Qy 822 AATGAAGGATAGAAATCCAGAGTAAGT 852  
Db 706 AATGAAGGATAGAAATCCAGAGTAAGT 736

## RESULT 15

US-10-113-872-1791  
; Sequence 1791, Application US/10113872  
; Publication No. US20030170255A1  
; GENERAL INFORMATION:  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Sleath, Paul R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Derrick  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C19  
; CURRENT APPLICATION NUMBER: US/10/113,872  
; CURRENT FILING DATE: 2002-03-28  
; NUMBER OF SEQ ID NOS: 2011  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-113-872-1791

Query Match 27.5%; Score 640; DB 16; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 162 GGAATCTCATCTTTCACGTGGTTTATGGTATTCATTCGTCGGCGCTCGACATCTGTA 221  
Db 46 GGAATCTCATCTTTCACGTGGTTTATGGTATTCATTCGTCGGCGCTCGACATCTGTA 105  
Qy 222 GCTGTCGTTTGGTTTGTATCTTTCATGATGATGATGATGATGATGATGATGATGATGAT 281  
Db 106 GCTGTCGTTTGGTTTGTATCTTTCATGATGATGATGATGATGATGATGATGATGATGAT 165  
Qy 282 GATGCTGATGATGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 166 GATGCTGATGATGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 225

```
Qy 342 GAGAGATCTACTTCAGAGCCAGCAGTCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401
Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 285
Qy 402 GAGGAGCAGGTTCTGTGGAGGCGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACAA 461
Db 286 GAGGAGCAGGTTCTGTGGAGGCGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACAA 345
Qy 462 ATTCAGTCCCTTCTCCATGAATGGTACACCCAGAACATGTTGAGGAGAGACTTGCAG 521
Db 346 ATTCAGTCCCTTCTCCATGAATGGTACACCCAGAACATGTTGAGGAGAGACTTGCAG 405
Qy 522 CAAGAAGATGACCCACAGGAGAACCAACAAAGAGATGATGAGTTCCTTATGGCGACT 581
Db 406 CAAGAAGATGACCCACAGGAGAACCAACAAAGAGATGATGAGTTCCTTATGGCGACT 465
Qy 582 GATGTAGATGATAGATTGAGACCCCTGGAACCTGAAAGTATCTCATGAAGAACCGAGCAT 641
Db 466 GATGTAGATGATAGATTGAGACCCCTGGAACCTGAAAGTATCTCATGAAGAACCGAGCAT 525
Qy 642 AGTTACCACTGGAGAGACAGTTCACAGACTGTAATCAGGATGATGAGAGATGATG 701
Db 526 AGTTACCACTGGAGAGACAGTTCACAGACTGTAATCAGGATGATGAGAGATGATG 585
Qy 702 TCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAGATGAAAGATTGCACCAT 761
Db 586 TCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAGATGAAAGATTGCACCAT 645
Qy 762 GATACAGATGATGTAACATACAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAA 821
Db 646 GATACAGATGATGTAACATACAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAA 705
Qy 822 AATGAGGGATAGAAATCACAGAACTAACTG 852
Db 706 AATGAGGGATAGAAATCACAGAACTAACTG 736
```

RESULT 16

```
US-10-283-017-1791
; Sequence 1791, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margaita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.47820
; CURRENT APPLICATION NUMBER: US/10/283,017
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1791
; LENGTH: 2442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-017-1791
Query Match 27.5%; Score 640; DB 17; Length 2442;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 690; Conservative 1; Mismatches 0; Gaps 0;
Qy 162 GGAACCTTCATCTTACGTGGTTTATGAGTTCATGCTGGCGCTGAGACATCTGTA 221
```

```
Db 46 GGAACCTTCATCTTACGTGGTTTATGAGTTCATGCTGGCGCTGAGACATCTGTA 105
Qy 222 GCTGTGCTTTGGTTTTCATCTTGTGTTGACTATGAGAAAGTTCCTAGGAAACTTAGGAATCTAT 281
Db 106 GCTGTGCTTTGGTTTTCATCTTGTGTTGACTATGAGAAAGTTCCTAGGAAACTTAGGAATCTAT 165
Qy 282 GATGCTGATGTTGATGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 341
Db 166 GATGCTGATGTTGATGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 225
Qy 342 GAGAGATCTACTTCAGAGCCAGCAGTCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401
Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 285
Qy 402 GAGGAGCAGGTTCTGTGGAGGCGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACAA 461
Db 286 GAGGAGCAGGTTCTGTGGAGGCGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAACAA 345
Qy 462 ATTCAGTCCCTTCTCCATGAATGGTACACCCAGAACATGTTGAGGAGAGACTTGCAG 521
Db 346 ATTCAGTCCCTTCTCCATGAATGGTACACCCAGAACATGTTGAGGAGAGACTTGCAG 405
Qy 522 CAAGAAGATGACCCACAGGAGAACCAACAAAGAGATGATGAGTTCCTTATGGCGACT 581
Db 406 CAAGAAGATGACCCACAGGAGAACCAACAAAGAGATGATGAGTTCCTTATGGCGACT 465
Qy 582 GATGTAGATGATAGATTGAGACCCCTGGAACCTGAAAGTATCTCATGAAGAACCGAGCAT 641
Db 466 GATGTAGATGATAGATTGAGACCCCTGGAACCTGAAAGTATCTCATGAAGAACCGAGCAT 525
Qy 642 AGTTACCACTGGAGAGACAGTTCACAGACTGTAATCAGGATGATGAGAGATGATG 701
Db 526 AGTTACCACTGGAGAGACAGTTCACAGACTGTAATCAGGATGATGAGAGATGATG 585
Qy 702 TCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAGATGAAAGATTGCACCAT 761
Db 586 TCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAGATGAAAGATTGCACCAT 645
Qy 762 GATACAGATGATGTAACATACAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAA 821
Db 646 GATACAGATGATGTAACATACAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAA 705
Qy 822 AATGAGGGATAGAAATCACAGAACTAACTG 852
Db 706 AATGAGGGATAGAAATCACAGAACTAACTG 736
```

RESULT 17

```
US-10-198-846-9938
; Sequence 9938, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9938
; LENGTH: 3110
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

/ LOCATION: 3009, 3010, 3011, 3012, 3013, 3014, 3015, 3016, 3017, 3018,  
/ LOCATION: 3019, 3020, 3021, 3022, 3023, 3024, 3025, 3026, 3027, 3028,  
/ LOCATION: 3029, 3030, 3031, 3032, 3033, 3034, 3035, 3036, 3037, 3038,  
/ LOCATION: 3039, 3040, 3041, 3042, 3043, 3044, 3045, 3046, 3047  
/ OTHER INFORMATION: n = A,T,C or G  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 3048, 3049, 3050, 3051, 3052, 3053, 3106, 3107, 3108, 3109,  
/ LOCATION: 3110  
/ OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-9938

Query Match 26.5%; Score 615; DB 14; Length 3110;  
Best Local Similarity 100.0%; Pred. No. 2.8e-307;  
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	332	AGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA	391
Db	595	AGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA	654
Qy	392	CAGTGGCCCGAGAGCAGGTTCTGTGGAGGAGAGAACCCAGAGATATCGAAGTGAAGC	451
Db	655	CAGTGGCCCGAGAGCAGGTTCTGTGGAGGAGAGAACCCAGAGATATCGAAGTGAAGC	714
Qy	452	AAAGAGCAAAATTCAGTCCCTTCTCCATGAAATGGTACACGAGAGATGTGGAGGAGA	511
Db	715	AAAGAGCAAAATTCAGTCCCTTCTCCATGAAATGGTACACGAGAGATGTGGAGGAGA	774
Qy	512	AGACTTGCACCAAGAGATGAGCCACAGAGAGAACCAACAGAGAGATGATGATTTCT	571
Db	775	AGACTTGCACCAAGAGATGAGCCACAGAGAGAACCAACAGAGAGATGATGATTTCT	834
Qy	572	TATGGGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	631
Db	835	TATGGGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	894
Qy	632	AACGAGCATGTTACACGCTGGAAGAGACAGTTTCAACAGCTGTAATCAGGATATGA	691
Db	895	AACGAGCATGTTACACGCTGGAAGAGACAGTTTCAACAGCTGTAATCAGGATATGA	954
Qy	692	AGAGATGATGTTCTGAGCAGGAGAAATCCAGATTCAGTGAACAGTAGTAGAAGATGAA	751
Db	955	AGAGATGATGTTCTGAGCAGGAGAAATCCAGATTCAGTGAACAGTAGTAGAAGATGAA	1014
Qy	752	ATTGCACCATGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATG	811
Db	1015	ATTGCACCATGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATG	1074
Qy	812	ACCTTAGAAGATGAGGATGAGAAATCAGAGATTAATCTCTCCCTGAGGATATCC	871
Db	1075	ACCTTAGAAGATGAGGATGAGAAATCAGAGATTAATCTCTCCCTGAGGATATCC	1134
Qy	872	TGTAGAAGATTCACAGTAAATTTAGAGAAGTAAAGCAATTTTCTGTGGAAGAACAGCA	931
Db	1135	TGTAGAAGATTCACAGTAAATTTAGAGAAGTAAAGCAATTTTCTGTGGAAGAACAGCA	1194
Qy	932	GGAGTACACACAGA 946	
Db	1195	GGAGTACACACAGA 1209	

RESULT 18  
US-09-918-995-34279  
; Sequence 34279, Application US/09918995  
; Publication No.: US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; OTHER INFORMATION: n = A,T,C or G

/ PRIOR FILING DATE: 1999-01-20  
/ NUMBER OF SEQ ID NOS: 38054  
/ SOFTWARE: FastSeq for Windows Version 3.0  
/ SEQ ID NO 34279  
/ LENGTH: 502  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: (1)...(502)  
/ OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-34279

Query Match	19.8%;	Score 459;	DB 10;	Length 502;
Best Local Similarity	100.0%;	Pred. No. 1.7e-226;		
Matches 459;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	912	TTTCTGTGGAAGAACAGCAGGAGGATACCCAGAGAAACAAATAGAAAAACAGATGATCCA	971	
Db	44	TTTCTGTGGAAGAACAGCAGGAGGATACCCAGAGAAACAAATAGAAAAACAGATGATCCA	103	
Qy	972	GAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATAAAATTTGATAAGACT	1031	
Db	104	GAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATAAAATTTGATAAGACT	163	
Qy	1032	ATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCGTAAAGGGGAAAAATTTGAGGAAGCA	1091	
Db	164	ATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCGTAAAGGGGAAAAATTTGAGGAAGCA	223	
Qy	1092	GTGATGCANTTTAAGAACTAGTACGCAATACCTCAGATCCACGAGCAAGATATGG	1151	
Db	224	GTGATGCANTTTAAGAACTAGTACGCAATACCTCAGATCCACGAGCAAGATATGG	283	
Qy	1152	AAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGATAGGTGCTACGTGA	1211	
Db	284	AAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGATAGGTGCTACGTGA	343	
Qy	1212	GCCATCCAGACCTTACCAAGAGGTGGCCAGCCTTACCTGATGTCCTGAGACCTGCTGAAG	1271	
Db	344	GCCATCCAGACCTTACCAAGAGGTGGCCAGCCTTACCTGATGTCCTGAGACCTGCTGAAG	403	
Qy	1272	CTGAGTTTGAAGCGTCCCTCAGACAGGCAACAATTTCTAGGTATATGAGAGGTTCCCTG	1331	
Db	404	CTGAGTTTGAAGCGTCCCTCAGACAGGCAACAATTTCTAGGTATATGAGAGGTTCCCTG	463	
Qy	1332	CTTACCTCCAGAGATAGTTCACACTATTTCCCAATGAT 1370		
Db	464	CTTACCTCCAGAGATAGTTCACACTATTTCCCAATGAT 502		

RESULT 19  
US-09-969-034-3202/c  
; Sequence 3202, Application US/09969034  
; Publication No. US20040110668A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Asule, Jon H.  
; APPLICANT: Carroll, Eddie III  
; APPLICANT: Catino, Theodore J.  
; APPLICANT: Dwivedi, Poornima  
; APPLICANT: Molino, Gary A.  
; APPLICANT: Thiagalingam, Arunthathi  
; APPLICANT: Lewis, Marcia B.  
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially  
; FILE REFERENCE: 1657/1032  
; CURRENT APPLICATION NUMBER: US/09/969,034  
; PRIOR FILING DATE: 2001-10-02  
; CURRENT APPLICATION NUMBER: 60/237,271  
; PRIOR FILING DATE: 2000-02-10  
; NUMBER OF SEQ ID NOS: 4494  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3202

	Query Match	13.6%	Score 315;	DB 17;	Length 366;
	Best Local Similarity	99.7%;	Pred. No. 6.4e-152;		
	Matches 365;	Conservative	0; Mismatches 1;	Indels 0;	Gaps 0
Qy	191	GATTGCATTTGCTGGCGTC	TGGACATCTGTAGCTGTGCTGGTGGTTTGATCTTTGATCTGA	250	
Dd	1	GATTGCATTTGCTGGCGTC	TGGACATCTGTAGCTGTGCTGGTGGTTTGATCTTTGATCTGA	60	
Qy	251	TGAGGAAGTTCTPAGAAAAC	CTTAGGAATCTTATGATCTGATGCTGATGGAGATTTTGATCT	310	
Dd	61	TGAGGAAGTTCTPAGAAAAC	CTTAGGAATCTTATGATCTGATGCTGATGGAGATTTTGATCT	120	
Qy	311	GGATGATGCCAAAGTTTTT	TATTAGGACTTAAGAAGAGATCTTACTTTCAGAGCCAGCAGTCCC	370	
Dd	121	GGATGATGCCAAAGTTTTT	TATTAGGACTTAAGAAGAGATCTTACTTTCAGAGCCAGCAGTCCC	180	
Qy	371	GCCAGAAGAGGCTGAGCCA	CACATGTAGCCCGGAGCAGGTTCTCTGTGGAGGACGCAACC	430	
Dd	181	GCCAGAAGAGGCTGAGCCA	CACATGTAGCCCGGAGCAGGTTCTCTGTGGAGGACGCAACC	240	
Qy	431	CCAGATATTCGAAGATGA	AGCAAAGAACAAATTCAGTCCCTTCTCCCATCAAAATGGTGACA	490	
Dd	241	CCAGATATTCGAAGATGA	AGCAAAGAACAAATTCAGTCCCTTCTCCCATCAAAATGGTGACA	300	

QY 491 CGCAGAACATGTTGAGGCGAGAACTTGCACCAAGAGATGGACCCACAGGAGAACCA 550  
| | | | |  
Db 301 CGCAGAACATGTTGAGGCGAGAACTTGCACCAAGAGATGGACCCACAGGAGAACCA 360  
| | | | |  
QY 551 ACAAGA 556  
| | | | |  
Db 361 ACAAGA 366

## RESULT 22

US-10-066-543-277  
; Sequence 277, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066.543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 277  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 341, 343  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-277

Query Match 10.7%; Score 249; DB 14; Length 495;  
Best Local Similarity 99.7%; Pred. No. 1e-117;  
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 630 GAAACCGAGCATAGTTACACGCTGGAGAGACAGTTTCACAGACTGTAATCAGGATG 689  
| | | | |  
Db 41 GAAACCGAGCATAGTTACACGCTGGAGAGACAGTTTCACAGACTGTAATCAGGATG 100  
| | | | |  
QY 690 GAAGAGATGATCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAAGTAA 749  
| | | | |  
Db 101 GAAGAGATGATCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAAGTAA 160  
| | | | |  
QY 750 AGATTGCCACCATGATACAGATGATGTAAATACCAAGTCTATGAGGAAACAAGCAGTATAT 809  
| | | | |  
Db 161 AGATTGCCACCATGATACAGATGATGTAAATACCAAGTCTATGAGGAAACAAGCAGTATAT 220  
| | | | |  
QY 810 GAACTCTAGAAATGAAGGATAGAAATACAGAGTAACTGCTCCCTGAGGATAT 869  
| | | | |  
Db 221 GAACTCTAGAAATGAAGGATAGAAATACAGAGTAACTGCTCCCTGAGGATAT 280  
| | | | |  
QY 870 CTTGTAGAGATTTCAGAGTAACTGTAGAGAGTAACTTTCTCTGTGGAAGACAG 929  
| | | | |  
Db 281 CTTGTAGAGATTTCAGAGTAACTGTAGAGAGTAACTTTCTCTGTGGAAGACAG 340  
| | | | |

## RESULT 23

US-10-242-535A-2707  
; Sequence 2707, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2707  
; LENGTH: 191  
; TYPE: DNA  
; ORGANISM: Human  
; US-10-242-535A-2707

Query Match 5.8%; Score 135; DB 17; Length 191;  
Best Local Similarity 100.0%; Pred. No. 1.1e-58;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 367 TCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCGAG 426  
| | | | |  
Db 1 TCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCGAG 60  
| | | | |  
QY 427 AACCCAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAATGG 486  
| | | | |  
Db 61 AACCCAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAATGG 120  
| | | | |  
QY 487 TACACGCGAGAACATG 501  
| | | | |  
Db 121 TACACGCGAGAACATG 135  
| | | | |

## RESULT 24

US-10-085-783A-2707  
; Sequence 2707, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2707  
; LENGTH: 191  
; TYPE: DNA  
; ORGANISM: Human  
; US-10-085-783A-2707

Query Match 5.8%; Score 135; DB 17; Length 191;  
Best Local Similarity 100.0%; Pred. No. 1.1e-58;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 367 TCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCGAG 426  
| | | | |  
Db 1 TCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCGAG 60  
| | | | |  
QY 427 AACCCAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAATGG 486  
| | | | |





```

; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2825
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-2825

```

```

Query Match      3.4%; Score 78; DB 17; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.7e-29;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      349  CTACTTCAGAGCCACAGTCCCGCCAGGAGGCTGAGCCACACTGAGCCCGAGGAGC 408
          |||||
          450  CTACTTCAGAGCCACAGTCCCGCCAGGAGGCTGAGCCACACTGAGCCCGAGGAGC 391

QY      409  AGGTTCTCTGGAGGCAG 426
          |||||
          db      390  AGGTTCTCTGGAGGCAG 373.

```

```

RESULT 29
US-10-443-622-24
, Sequence 24, Application US/10443622
, Publication No. US20040024192A1
, GENERAL INFORMATION:
, APPLICANT: Carter et al.
, TITLE OF INVENTION: 19 Human Secreted Proteins
, FILE REFERENCE: P2009p1
, CURRENT APPLICATION NUMBER: US/10/443,622
, CURRENT FILING DATE: 2003-05-23
, PRIOR APPLICATION NUMBER: PCT/US98/13608
, PRIOR FILING DATE: 1998-06-30
, PRIOR APPLICATION NUMBER: 60/051,480
, PRIOR FILING DATE: 1997-07-01
, PRIOR APPLICATION NUMBER: 60/051,381
, PRIOR FILING DATE: 1997-07-01
, PRIOR APPLICATION NUMBER: 60/058,663
, PRIOR FILING DATE: 1997-09-12
, PRIOR APPLICATION NUMBER: 60/058,598
, PRIOR FILING DATE: 1997-09-12
, NUMBER OF SEQ ID NOS: 156
, SOFTWARE: Patentin Ver. 2.0
, SEQ ID NO 24
, LENGTH: 1382
, TYPE: DNA
, ORGANISM: Homo sapiens
US-10-443-622-24

```

Query Match	3.1%	Score 72	DB 17	Length 1382
Best Local Similarity	100.0%	Pred. No. 5e-26		
Matches 72	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	263	AGGAAACTAGGAATCTATGATCTGATCGTGTGATGAGATTTTGGATGTGATGATGCCAA	322	
Db	233	AGGAAACTAGGAATCTATGATCTGATCGTGTGATGAGATTTTGGATGTGATGATGCCAA	292	
QY	323	AGTTTATTATTAGG	334	

```

Db      293   AGTTTATTAGG 304
|||||
RESULT 30
US-908-975-11070
Sequence 11070, Application US-
Publication No. US20030165943A2
GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: WASSERMAN, Alon
APPLICANT: MINTZ, Eli
APPLICANT: MINTZ, Liat
APPLICANT: FAGLER, Simchon
TITLE OF INVENTION: OLIGONUCLEOTIDE SEQUENCES AND METHODS FOR IDENTIFYING AND ANALYZING DNA
TITLE OF INVENTION: THAT POP
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US-908-975-11070
PRIOR FILING DATE: 2001-07-28
PRIOR APPLICATION NUMBER: US-908-975-11070
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US-908-975-11070
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.1.1
SEQ ID NO 11070
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-908-975-11070

```

Query Match 2.6%; Score 60; DB 10; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-20;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 GACTGTAATCAGGATATGGAAGAGATGATGTCGAGCAGGAAAAATCCAGATTCACAGTAA 731  
 Db 1 GACTGTAATCAGGATATGGAAGAGATGATGTCGAGCAGGAAAAATCCAGATTCACAGTAA 60

```

RESULT 31
US-10-116-275-322/c
/ Sequence 322, Application US/10116275
/ Publication No. US20030211476A1
/ GENERAL INFORMATION:
/ APPLICANT: Elian Pharmaceutical Technology
/ APPLICANT: O'Mahony, Daniel J.
/ APPLICANT: Brayden, David
/ APPLICANT: Byrne, Daragh
/ APPLICANT: Lambkin, Imelda
/ APPLICANT: Higgins, Lisa
/ TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
/ FILE REFERENCE: E1067/20087
/ CURRENT APPLICATION NUMBER: US/10/116,275
/ CURRENT FILING DATE: 2002-10-04
/ NUMBER OF SEQ ID NOS: 349
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 322
/ LENGTH: 2208
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-116-275-322

```

Query Match 2.5%; Score 59; DB 17; Length 2208;  
Best Local Similarity 100.0%; Pred. No. 2.8e-13;  
Matches: 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
RESULT 32
US-10-755-889-177/c
; Sequence 177, Application US/10755889
; Publication No. US20040171822A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 177
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-177

Query Match      2.5%; Score 59; DB 18; Length 2208;
Best Local Similarity 100.0%; Pred. No. 2.8e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2239 GGCATCCGGAAGTGCACACGAGAGACGCGCTTCCAGCAATTTAGCATGAATTC 2297
Db 59 GGCATCCGGAAGTGCACACGAGAGACGCGCTTCCAGCAATTTAGCATGAATTC 1

RESULT 33
US-09-969-034-3266/c
; Sequence 3266, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3266
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 503, 512, 553, 565, 570, 577, 578
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3266

Query Match      2.2%; Score 51; DB 11; Length 583;
Best Local Similarity 100.0%; Pred. No. 3.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2274 CTTCCAGCAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTCGAGAGA 2324
Db 439 CTTCCAGCAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTCGAGAGA 389
```

```
RESULT 34
US-10-027-632-7020
; Sequence 7020, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7020
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-7020

Query Match      1.9%; Score 44; DB 13; Length 591;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 800 AGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCACAG 843
Db 304 AGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCACAG 347

RESULT 35
US-10-027-632-7020
; Sequence 7020, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7020
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Human
```

## US-10-027-632-7020

Query Match 1.7% Score 44; DB 17; Length 591;  
Best Local Similarity 100.0%; Pred. No. 1.5e-11; Length 591;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 800 ACCAGTATATGAACTCTAGAAAATGAAGGATAGAAATCACAG 843  
Db 304 ACCAGTATATGAACTCTAGAAAATGAAGGATAGAAATCACAG 347

## RESULT 36

US-10-027-632-56205/c  
; Sequence 56205, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720-  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56205  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-56205

Query Match 1.7% Score 39; DB 13; Length 453;  
Best Local Similarity 100.0%; Pred. No. 6e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 387  
Db 453 CTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 415

## RESULT 37

US-10-027-632-310460/c  
; Sequence 310460, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 310460  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-310460

Query Match 1.7% Score 39; DB 13; Length 453;  
Best Local Similarity 100.0%; Pred. No. 6e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 387  
Db 453 CTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 415

## RESULT 38

US-10-027-632-56205/c  
; Sequence 56205, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56205  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-56205

Query Match 1.7% Score 39; DB 17; Length 453;  
Best Local Similarity 100.0%; Pred. No. 6e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 387  
Db 453 CTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 415

## RESULT 39

US-10-027-632-310460/c  
; Sequence 310460, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 310460  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Human  
; US-10-632-310460

Query Match 1.7%; Score 39; DB 17; Length 453;  
Best Local Similarity 100.0%; Pred. No. 6e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CTACTTCAGCCGACGAGTCCCGCCAGAGAGGCTGAGC 387  
Db 453 CTACTTCAGCCGACGAGTCCCGCCAGAGAGGCTGAGC 415

## RESULT 40

US-10-424-599-97119  
; Sequence 97119, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 97119  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_58710C.1  
US-10-424-599-97119

Query Match 1.0%; Score 24; DB 17; Length 859;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 AGCAGCGCAACAGCAGCAGCAGC 59  
Db 128 AGCAGCGCAACAGCAGCAGCAGC 151

## RESULT 41

US-10-425-115-124491/c  
; Sequence 124491, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 124491  
; LENGTH: 2135  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_45013C.1  
US-10-425-115-124491

Query Match 1.0%; Score 23; DB 18; Length 2135;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 GCAGCGGCAACAGCAGCAGCAGC 59  
Db 1991 GCAGCGGCAACAGCAGCAGCAGC 1969

## RESULT 42

US-10-437-963-73067  
; Sequence 73067, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 73067  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_73384C.1  
US-10-437-963-73067

Query Match 0.9%; Score 21; DB 18; Length 681;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CGGCAACAGCAGCAGCAGCAGC 61  
Db 39 CGGCAACAGCAGCAGCAGCAGC 59

## RESULT 43

US-10-425-115-42949/c  
; Sequence 42949, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 42949
; LENGTH: 715
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577.139178C.1
US-10-425-115-42949

Query Match      0.9%; Score 21; DB 18; Length 715;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GCGGCAACAGCAGCAGCG 60
    |||||
Db 462 GCGGCAACAGCAGCAGCG 442

RESULT 44
US-10-308-503-188
; Sequence 188, Application US/10308503
; Publication No. US20030191080A1
; GENERAL INFORMATION:
; APPLICANT: PHILLIPS, M. IAN
; TITLE OF INVENTION: ANTISENSE COMPOSITIONS TARGETED TO BETAL-ADRENOCEPTOR-SPECIFIC mRNA
; FILE REFERENCE: 4300.013900
; CURRENT APPLICATION NUMBER: US/10/308,503
; PRIOR FILING DATE: 2003-12-03
; PRIOR APPLICATION NUMBER: US/09/614,034
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 09/152,717
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: PCT/US99/21007
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 188
; TYPE: DNA
; LENGTH: 1845
; ORGANISM: UNKNOWN
; FEATURE:
; OTHER INFORMATION: SYNTHETIC OLIGONUCLEOTIDE
; NAME/KEY: misc_feature
; LOCATION: (1828)..(1828)
; OTHER INFORMATION: WHERE N = A, T, C OR G
US-10-308-503-188

Query Match      0.9%; Score 21; DB 16; Length 1845;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CAGCAGCGGCTCCGGCAGCG 73
    |||||
Db 93 CAGCAGCGGCTCCGGCAGCG 113

RESULT 45
US-09-859-604-10/c
; Sequence 10, Application US/09859604
; Patent No. US20020110559A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M
; APPLICANT: Deutch, Alan H
; APPLICANT: Ghanbari, Hossein A
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 CIP
; CURRENT APPLICATION NUMBER: US/09/859,604
; PRIOR FILING DATE: 2001-05-17
; OTHER INFORMATION:
; NUMBER OF SEQ ID NOS: 13
```

```
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Location (-1)
; OTHER INFORMATION: oligonucleotide
US-09-859-604-10

Query Match      0.9%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AATGCCCCAGCGTAAAGTG 30
    |||||
Db 20 AATGCCCCAGCGTAAAGTG 1

RESULT 46
US-09-859-604-11/c
; Sequence 11, Application US/09859604
; Patent No. US20020110559A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M
; APPLICANT: Deutch, Alan H
; APPLICANT: Ghanbari, Hossein A
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 CIP
; CURRENT APPLICATION NUMBER: US/09/859,604
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; TYPE: DNA
; LENGTH: 20
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Location (-6)
; OTHER INFORMATION: oligonucleotide
US-09-859-604-11

Query Match      0.9%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CGTGAATGGCCAGCGTAA 25
    |||||
Db 20 CGTGAATGGCCAGCGTAA 1

RESULT 47
US-09-859-604-12/c
; Sequence 12, Application US/09859604
; Patent No. US20020110559A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M
; APPLICANT: Deutch, Alan H
; APPLICANT: Ghanbari, Hossein A
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 CIP
; CURRENT APPLICATION NUMBER: US/09/859,604
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 13
```

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 12  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Locations (-11)  
; OTHER INFORMATION: oligonucleotide  
US-09-859-604-12

Query Match 0.9%; Score 20; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCATGGCCAG 20  
Db 20 CGGACCGTGCATGGCCAG 1

RESULT 48

US-10-425-115-21348/c  
; Sequence 21348, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 21348  
; LENGTH: 183  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MFT4577\_119464C.1  
US-10-425-115-21348

Query Match 0.9%; Score 20; DB 18; Length 183;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 CGAGTGGCGGCGAGCAGC 98  
Db 59 CGAGTGGCGGCGAGCAGC 40

RESULT 49

US-10-425-115-22449  
; Sequence 22449, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 22449  
; LENGTH: 432  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MFT4577\_120475C.1  
US-10-425-115-22449

Query Match 0.9%; Score 20; DB 18; Length 432;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GCGGCAACAGCAGCAGC 59  
Db 396 GCGGCAACAGCAGCAGC 415

RESULT 50

US-10-767-701-26002/c  
; Sequence 26002, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 26002  
; LENGTH: 489  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 30969203  
US-10-767-701-26002

Query Match 0.9%; Score 20; DB 18; Length 489;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 AGCAGCAGCGGCTCCGCGAG 70  
Db 242 AGCAGCAGCGGCTCCGCGAG 223

RESULT 51

US-10-424-599-89196  
; Sequence 89196, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 89196  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(495)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_51552C.1  
US-10-424-599-89196

Query Match 0.9%; Score 20; DB 17; Length 495;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 991 AGAAAAAGAGCCCTAACTT 1010

```

Db      402 AGAAGAGAGCCCTAAACTT 421
|||||
Query Match      0.9%; Score 20; DB 13; Length 567;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 52
US-10-029-386-9144/c
; Sequence 9144, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9144
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007262.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: SWISSPROT HIT: P54258, EVALUATE 1.80e+00
; OTHER INFORMATION: NT HIT: AB051488.1, EVALUATE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: B1458804.1, EVALUATE 7.00e-63
US-10-029-386-9144

Query Match      0.9%; Score 20; DB 16; Length 550;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1256 TGCAGACCTGCTGAAGCTGA 1275
|||||
Db      280 TGCAGACCTGCTGAAGCTGA 261

RESULT 53
US-10-027-632-82527/c
; Sequence 82527, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82527
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-82527

Query Match      0.9%; Score 20; DB 13; Length 567;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      972 GAACAAAAAGCAAAAGTTAA 991
|||||
Db      212 GAACAAAAAGCAAAAGTTAA 193

RESULT 54
US-10-027-632-322022/c
; Sequence 322022, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322022
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-322022

Query Match      0.9%; Score 20; DB 13; Length 567;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      972 GAACAAAAAGCAAAAGTTAA 991
|||||
Db      212 GAACAAAAAGCAAAAGTTAA 193

RESULT 55
US-10-027-632-322023/c
; Sequence 322023, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24

```



; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 322023  
; LENGTH: 567  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-322023

Query Match 0.9%; Score 20; DB 13; Length 567;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
Db 212 GAACAAAAGCAAAAGTTAA 193

RESULT 56  
US-10-027-632-82527/c  
; Sequence 82527, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 82527  
; LENGTH: 567  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-82527

Query Match 0.9%; Score 20; DB 17; Length 567;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
Db 212 GAACAAAAGCAAAAGTTAA 193

RESULT 57  
US-10-027-632-322022/c  
; Sequence 322022, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 322023  
; LENGTH: 567  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-322023

Query Match 0.9%; Score 20; DB 17; Length 567;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
Db 212 GAACAAAAGCAAAAGTTAA 193

; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 322022  
; LENGTH: 567  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-322022

Query Match 0.9%; Score 20; DB 17; Length 567;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
Db 212 GAACAAAAGCAAAAGTTAA 193

RESULT 58  
US-10-027-632-322023/c  
; Sequence 322023, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 322023  
; LENGTH: 567  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-322023

Query Match 0.9%; Score 20; DB 17; Length 567;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
Db 212 GAACAAAAGCAAAAGTTAA 193

Db 212 GAACAAAAGCAAAAGTTAA 193

## RESULT 59

US-10-425-115-112661/c  
; Sequence 112661, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: LI, GANG  
; TITLE OF INVENTION: RAAV VECTOR-BASED PRO-OPIOMELANOCORTIN COMPOSITIONS AND METHODS  
; FILE REFERENCE: 4300.015400  
; CURRENT APPLICATION NUMBER: US/10/822,613  
; CURRENT FILING DATE: 2004-04-12  
; PRIOR FILING DATE: 2003-04-11  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 804  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-822-613-1  
Query Match 0.9%; Score 20; DB 18; Length 804;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-10-425-115-112661

Query Match 0.9%; Score 20; DB 18; Length 750;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1933 GCAAGGAGCTCCTAAACC 1952  
Db 412 GCAAGGAGCTCCTAAACC 393

## RESULT 60

US-10-437-963-71261  
; Sequence 71261, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 71261  
; LENGTH: 771  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_71751C.1  
US-10-437-963-71261

Query Match 0.9%; Score 20; DB 18; Length 771;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCGCGC 62  
Db 372 GCAACAGCAGCAGCGCGC 391

## RESULT 61

US-10-822-613-1

; Sequence 1, Application US/10822613  
; Publication No. US20050002905A1  
; GENERAL INFORMATION:  
; APPLICANT: SCARPACE, PHILIP J.

; APPLICANT: LI, GANG  
; TITLE OF INVENTION: RAAV VECTOR-BASED PRO-OPIOMELANOCORTIN COMPOSITIONS AND METHODS  
; FILE REFERENCE: 4300.015400  
; CURRENT APPLICATION NUMBER: US/10/822,613  
; CURRENT FILING DATE: 2004-04-12  
; PRIOR FILING DATE: 2003-04-11  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 804  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-822-613-1

Query Match 0.9%; Score 20; DB 18; Length 804;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCGCGC 62  
Db 269 GCAACAGCAGCAGCGCGC 288

## RESULT 62

US-10-437-963-24406  
; Sequence 24406, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 24406  
; LENGTH: 843  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_29392C.1  
US-10-437-963-24406

Query Match 0.9%; Score 20; DB 18; Length 843;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 AGCAGCGCTCCGCGCGCG 73  
Db 366 AGCAGCGCTCCGCGCGCG 385

## RESULT 63

US-10-437-963-51503  
; Sequence 51503, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 51503
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53889C.1
US-10-437-963-51503

Query Match          0.9%; Score 20; DB 18; Length 924;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GCGCAACAGCAGCAGCAGC 59
Db 663 GCGCAACAGCAGCAGCAGC 682

RESULT 64
US-10-646-620A-8
; Sequence 8, Application US/10646620A
; Publication No. US20040138115A1
; GENERAL INFORMATION:
; APPLICANT: WITTIG, Burghardt
; APPLICANT: STEIN, Christoph
; APPLICANT: SCHAEFER, Michael
; APPLICANT: SCHROFF, Matthias
; APPLICANT: JUNGHANS, Claas
; APPLICANT: KOENIG MEREDIZ, Sven A.
; TITLE OF INVENTION: Local Pain Combating-Agent
; FILE REFERENCE: NHL-NP-43
; CURRENT APPLICATION NUMBER: US/10/646,620A
; CURRENT FILING DATE: 2003-08-22
; PRIOR FILING DATE: 2001-09-09
; PRIOR FILING DATE: 2001-02-24
; PRIOR FILING DATE: PCT/DE02/00583
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 8
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (847)..(933)
; OTHER INFORMATION: beta-endorphin cDNA sequence
US-10-646-620A-8

Query Match          0.9%; Score 20; DB 18; Length 936;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62
Db 401 GCAACAGCAGCAGCAGCGGC 420

RESULT 65
US-10-425-115-180542/c
; Sequence 180542, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: P2027P1
; CURRENT APPLICATION NUMBER: US/09/397,945
```

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180542
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1011)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_96234C.1
US-10-425-115-180542

Query Match          0.9%; Score 20; DB 18; Length 1011;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 AGCAGCAGCGCTCCGGCAG 70
Db 229 AGCAGCAGCGCTCCGGCAG 210

RESULT 66
US-09-853-386-105
; Sequence 105, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Bresnihan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; FILE REFERENCE: P01972US1
; CURRENT APPLICATION NUMBER: US/09/853,386
; CURRENT FILING DATE: 2001-05-11
; PRIOR FILING DATE: 2001-05-11
; PRIOR FILING DATE: US 60/203645
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 105
; LENGTH: 1072
; TYPE: DNA
; ORGANISM: HUMAN
US-09-853-386-105

Query Match          0.9%; Score 20; DB 9; Length 1072;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62
Db 375 GCAACAGCAGCAGCAGCGGC 394

RESULT 67
US-09-397-945-80/c
; Sequence 80, Application US/09397945
; Publication No. US20030065139A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: P2027P1
; CURRENT APPLICATION NUMBER: US/09/397,945
```

; CURRENT FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: PCT/US99/05804  
; PRIOR FILING DATE: 1999-03-18  
; PRIOR APPLICATION NUMBER: 60/078,566  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,576  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,573  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,574  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,579  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/080,314  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080,312  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/078,578  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,581  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,577  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,563  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/080,313  
; PRIOR FILING DATE: 1998-04-01  
; NUMBER OF SEQ ID NOS: 470  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 80  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-397-945-80

Query Match 0.9%; Score 20; DB 10; Length 1109;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62  
DB 152 GCAACAGCAGCAGCAGCGGC 133

RESULT 68  
US-10-653-595-80/c  
; Sequence 80, Application US/10653595  
; Publication No. US20040048304A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et. al.  
; TITLE OF INVENTION: Human secreted proteins  
; FILE REFERENCE: P2027P1C  
; CURRENT APPLICATION NUMBER: US/10/653,595  
; CURRENT FILING DATE: 2003-09-03/397945  
; PRIOR APPLICATION NUMBER: US 09/397945  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: PCT/US99/05804  
; PRIOR FILING DATE: 1999-03-18  
; PRIOR APPLICATION NUMBER: 60/078,566  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,576  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,573  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,574  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/078,579  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/080,314  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080,312  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/078,578

; PRIOR FILING DATE: 1998-03-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 470  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 80  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-653-595-80

Query Match 0.9%; Score 20; DB 17; Length 1109;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62  
DB 152 GCAACAGCAGCAGCAGCGGC 133

RESULT 69  
US-09-960-288-1  
; Sequence 1, Application US/09960288  
; Patent No. US20020064833A1  
; GENERAL INFORMATION:  
; APPLICANT: Fagerlund, Tore  
; APPLICANT: Alestrom, Peter  
; APPLICANT: Berg, K+Fe  
; TITLE OF INVENTION: Recombinant plasmids and method for treating substance abuse  
; FILE REFERENCE: 406801  
; CURRENT APPLICATION NUMBER: US/09/960,288  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/426,877  
; PRIOR FILING DATE: 1999-10-26  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-960-288-1

Query Match 0.9%; Score 20; DB 9; Length 1230;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62  
DB 474 GCAACAGCAGCAGCAGCGGC 493

RESULT 70  
US-10-287-971-37/c  
; Sequence 37, Application US/10287971  
; Publication No. US20040067882A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook, et al  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-480A  
; CURRENT APPLICATION NUMBER: US/10/287,971  
; CURRENT FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 09/997,425  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: 10/035,568  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: 60/338,626  
; PRIOR FILING DATE: 2001-11-05  
; PRIOR APPLICATION NUMBER: 60/401,479  
; PRIOR FILING DATE: 2002-08-06  
; PRIOR APPLICATION NUMBER: 60/333,072  
; PRIOR FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: 60/348,283  
; PRIOR FILING DATE: 2001-11-09  
; PRIOR APPLICATION NUMBER: 60/393,262

; PRIOR FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/406,181  
; PRIOR FILING DATE: 2002-08-26  
; NUMBER OF SEQ ID NOS: 397  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 37  
; LENGTH: 1245  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (90)..(1091)  
US-10-287-971-37

Query Match 0.9%; Score 20; DB 17; Length 1245;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
Db 138 GCAACAGCAGCAGCAGCGGC 119

RESULT 71  
US-10-425-114-23644  
; Sequence 23644, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 23644  
; LENGTH: 1282  
; TYPE: DNA  
; ORGANISM: Oryza sativa nipponbare  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3599-031-D8\_FLI  
US-10-425-114-23644

Query Match 0.9%; Score 20; DB 17; Length 1282;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GCGGCAACAGCAGCAGCAGC 59  
Db 562 GCGGCAACAGCAGCAGCAGC 581

RESULT 72  
US-10-425-115-19062  
; Sequence 19062, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 19062

; LENGTH: 1435  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_117387C.1  
US-10-425-115-19062

Query Match 0.9%; Score 20; DB 18; Length 1435;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1908 CAACGAAGAAATGAAA 1927  
Db 1218 CAACGAAGAAATGAAA 1237

RESULT 73  
US-10-913-196-2/c  
; Sequence 2, Application US/10913196  
; Publication No. US2005053988A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
; APPLICANT: HUMAN SERVICES  
; APPLICANT: Pastan, Ira H.  
; APPLICANT: Eglund, Kristi  
; APPLICANT: Lee, Byungkook  
; APPLICANT: Vincent, James  
; TITLE OF INVENTION: GENE EXPRESSED IN BREAST CANCER AND METHODS OF USE  
; FILE REFERENCE: 4239-66345-02  
; CURRENT APPLICATION NUMBER: US/10/913,196  
; CURRENT FILING DATE: 2004-08-05  
; PRIOR APPLICATION NUMBER: US 60/493,522  
; PRIOR FILING DATE: 2003-08-08  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 1439  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-913-196-2

Query Match 0.9%; Score 20; DB 19; Length 1439;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
Db 483 GCAACAGCAGCAGCAGCGGC 464

RESULT 74  
US-10-425-114-34740/c  
; Sequence 34740, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 34740  
; LENGTH: 1850  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:

OTHER INFORMATION: Clone ID: UC-ZMFLMO17307H04\_FLI  
US-10-425-114-34740  
Query Match 0.9%; Score 20; DB 17; Length 1850;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 47 CAGCAGCAGCAGCGGCTCCG 66  
DB 485 CAGCAGCAGCAGCGGCTCCG 466  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (193)...(1380)  
OTHER INFORMATION:  
US-10-310-154-280  
Query Match 0.9%; Score 20; DB 17; Length 1891;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 40 GCGGCAACACGACGACGACG 59  
DB 1125 GCGGCAACACGACGACGACG 1144  
Search completed: March 26, 2005, 01:14:24  
Job time : 1240 secs

OTHER INFORMATION: Clone ID: UC-ZMFLMO17307H04\_FLI  
US-10-425-114-34740  
Query Match 0.9%; Score 20; DB 17; Length 1850;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 47 CAGCAGCAGCAGCGGCTCCG 66  
DB 485 CAGCAGCAGCAGCGGCTCCG 466  
RESULT 75  
US-10-310-154-280  
Sequence 280 Application US/10310154  
Publication No: US20030233670A1  
GENERAL INFORMATION:  
APPLICANT: Edgerton, Michael D  
APPLICANT: Chomet, Paul S.  
APPLICANT: Adams, Thomas H  
APPLICANT: Ruff, Thomas G.  
APPLICANT: Agarwal, Ameeta K.  
APPLICANT: Ahrens, Jeffrey E.  
APPLICANT: Ball, James A.  
APPLICANT: Banu, G.  
APPLICANT: Bell, Erin  
APPLICANT: Boddupalli, Raghava  
APPLICANT: Deikman, Jill  
APPLICANT: Deng, Molian  
APPLICANT: Dong, Jinzhao  
APPLICANT: Duff, Stephen M.  
APPLICANT: Galligan, Meghan M.  
APPLICANT: Hinchey, Brenda S.  
APPLICANT: Huang, Shihshieh  
APPLICANT: Johnson, G. Richard  
APPLICANT: Jung, Vincent  
APPLICANT: Kretzmer, Keith A  
APPLICANT: Laccetti, Lucille B.  
APPLICANT: Lai, Chao-Qiang  
APPLICANT: Lee, Gary  
APPLICANT: Lin, Jier-Yi  
APPLICANT: Liu, Jingdong  
APPLICANT: Lu, Bin  
APPLICANT: Luethy, Michael M.  
APPLICANT: Lund, Adrian  
APPLICANT: Madson, Linda L.  
APPLICANT: Malloy, Kathleen A.  
APPLICANT: McKiel, Christine L.  
APPLICANT: Miller, Philip W.  
APPLICANT: Padmavathi, Manchikanti  
APPLICANT: Parnell, Laurence D.  
APPLICANT: Start, William G.  
APPLICANT: Tennesen, Dan  
APPLICANT: Vidya, K.R.  
APPLICANT: Wang, Haiyun  
APPLICANT: Xin, Zhanqiao  
APPLICANT: Xu, Nanfei  
APPLICANT: Yang, Chunzhi  
APPLICANT: Zeng, Xiaoping  
APPLICANT: Zhang, Qiang  
APPLICANT: Zhao, Yajuan  
APPLICANT: Zhou, Li  
TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants  
FILE REFERENCE: 38-15(52796)B  
CURRENT APPLICATION NUMBER: US/10/310,154  
CURRENT FILING DATE: 2002-12-04  
PRIOR APPLICATION NUMBER: 60/337,358  
PRIOR FILING DATE: 2001-12-04  
NUMBER OF SEQ ID NOS: 736  
SEQ ID NO 280  
LENGTH: 1891  
TYPE: DNA  
ORGANISM: Oryza sativa

**This Page Blank (uspto)**

11

12

13

14

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 17:12:50 ; Search time 408 Seconds  
(without alignments)  
9320.361 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	4	US-09-903-248-3
2	2324	100.0	2324	4	US-09-903-199-3
3	2324	100.0	2324	4	US-09-903-216-3
4	2324	100.0	2324	4	US-09-903-063-3
5	2324	100.0	2324	4	US-09-859-604-3
6	785	33.8	2442	3	US-09-040-485-1
7	640	27.5	2442	4	US-09-702-705-1791
8	640	27.5	2442	4	US-09-736-457-1791
9	640	27.5	2442	4	US-09-671-325-1791
10	640	27.5	2442	4	US-09-659-521-1
11	130	5.6	74730	4	US-09-948-016-15189
12	115	4.9	956	4	US-09-948-016-3447
13	88	3.8	5746	4	US-09-948-016-16453
14	77	3.3	601	4	US-09-948-016-122372
15	73	3.1	601	4	US-09-948-016-122387
16	59	2.5	2208	4	US-09-948-016-122387
17	45	1.9	601	4	US-09-948-016-236
18	22	0.9	9984	4	US-09-948-016-13283
19	21	0.9	1845	4	US-09-614-034-188
20	20	0.9	20	4	US-09-859-604-10
21	20	0.9	20	4	US-09-859-604-11
22	20	0.9	342	4	US-09-859-604-12
23	20	0.9	786	4	US-09-621-976-24
24	20	0.9	1933	4	US-09-489-039A-1933
25	20	0.9	1071	4	US-09-948-016-136
26	20	0.9	1106	4	US-09-948-016-3254
27	20	0.9	1116	4	US-09-948-016-1812
28	20	0.9	11668	4	US-09-949-016-14996
29	20	0.9	11713	4	US-09-949-016-11878
30	20	0.9	47284	4	US-09-949-016-13554
31	20	0.9	117937	4	US-09-949-016-17029
32	20	0.9	117937	4	US-09-949-016-12762
33	20	0.9	117937	4	US-09-949-016-15775
34	19	0.8	148	4	US-09-513-999C-35281
35	19	0.8	309	4	US-09-248-796A-13462
36	19	0.8	601	4	US-09-949-016-61822
37	19	0.8	951	4	US-09-248-796A-4406
38	19	0.8	1599	4	US-09-949-016-1792
39	19	0.8	1623	3	US-08-339-214-25
40	19	0.8	1751	3	US-09-148-476-110
41	19	0.8	1821	4	US-09-948-016-5583
42	19	0.8	2286	4	US-09-620-312D-49
43	19	0.8	2377	3	US-09-149-476-264
44	19	0.8	3462	4	US-09-949-016-6642
45	19	0.8	3482	4	US-09-949-016-633
46	19	0.8	6625	4	US-09-949-016-13534
47	19	0.8	7460	4	US-09-949-016-12375
48	19	0.8	7462	4	US-09-949-016-14384
49	19	0.8	7700	4	US-09-949-016-15619
50	19	0.8	28843	4	US-09-949-016-17325
51	19	0.8	300598	4	US-09-949-016-11868
52	19	0.8	302604	4	US-09-949-016-14588
53	19	0.8	302604	4	US-09-949-016-14589
54	19	0.8	308362	4	US-09-948-016-17119
55	19	0.8	828152	4	US-09-948-016-12777
56	18	0.8	22	4	US-09-859-604-13
57	18	0.8	286	4	US-09-313-294A-3024
58	18	0.8	319	2	US-08-482-842B-2
59	18	0.8	452	4	US-09-513-999C-10601
60	18	0.8	529	4	US-09-270-767-10408
61	18	0.8	557	4	US-09-918-686-4
62	18	0.8	561	4	US-09-621-976-16896
63	18	0.8	601	4	US-09-949-016-36982
64	18	0.8	601	4	US-09-949-016-40425
65	18	0.8	601	4	US-09-949-016-42135
66	18	0.8	601	4	US-09-949-016-42136
67	18	0.8	601	4	US-09-949-016-42137
68	18	0.8	601	4	US-09-949-016-42138
69	18	0.8	601	4	US-09-948-016-42139
70	18	0.8	601	4	US-09-948-016-42140
71	18	0.8	601	4	US-09-949-016-53139
72	18	0.8	601	4	US-09-949-016-67303
73	18	0.8	601	4	US-09-949-016-69417
74	18	0.8	601	4	US-09-949-016-69418
75	18	0.8	601	4	US-09-949-016-69419
76	18	0.8	601	4	US-09-949-016-79910
77	18	0.8	601	4	US-09-949-016-79911
78	18	0.8	601	4	US-09-949-016-79912
79	18	0.8	601	4	US-09-949-016-80517
80	18	0.8	601	4	US-09-948-016-113358
81	18	0.8	601	4	US-09-948-016-113359
82	18	0.8	601	4	US-09-948-016-113360
83	18	0.8	601	4	US-09-948-016-153469
84	18	0.8	601	4	US-09-948-016-176839
85	18	0.8	601	4	US-09-949-016-177100
86	18	0.8	601	4	US-09-949-016-177101
87	18	0.8	601	4	US-09-949-016-177102
88	18	0.8	603	4	US-09-621-976-3826
89	18	0.8	615	4	US-09-513-999C-3903
90	18	0.8	625	4	US-09-270-767-13169
91	18	0.8	711	4	US-09-252-991A-2925
92	18	0.8	905	4	US-09-270-767-12936
93	18	0.8	934	3	US-09-174-465D-4
94	18	0.8	934	3	US-09-599-564A-4
95	18	0.8	968	3	US-09-457-046B-13
96	18	0.8	968	3	US-09-866-570B-13
97	18	0.8	988	4	US-09-948-016-2721
98	18	0.8	1014	2	US-09-068-074-1
99	18	0.8	1014	2	US-08-555-912A-1
100	18	0.8	1014	4	US-09-275-900-1



## ALIGNMENTS

## RESULT 1

US-09-903-248-3  
; Sequence 3, Application US/09903248  
; Patent No. 6783758  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV5  
; CURRENT APPLICATION NUMBER: US/09/903,248  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-248-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGGACCTGCAATGGCCAGCGTAAGAAATGCCAAGACAGCGGCAACAGCAGCAGCAGCG	60
Db	1	CGGACCTGTGATGGCCAGCGTAAGAAATGCCAAGACAGCGGCAACAGCAGCAGCAGCG	60
Qy	61	GCTCCGCGAGCGTACGACGAGTGGCGGCGAGCAGCAGCGCGCGCGCGCGGAGACAA	120
Db	61	GCTCCGCGAGCGGTAGCAGCAGTGGCGGCGAGCAGCAGCGCGCGCGCGGAGACAA	120
Qy	121	AGCATGAGGACACAAAGATGGGAGGAAAGCGGACCTCGGGAACCTTCACTTCACGT	180
Db	121	AGCATGAGGACACAAAGATGGGAGGAAAGCGGACCTCGGGAACCTTCACTTCACGT	180
Qy	181	GGTTATGTTGATTCGATTCGCGGCGTCTGGACATCTGTAGCTGTGCTTTGTTGATC	240
Db	181	GGTTATGTTGATTCGATTCGCGGCGTCTGGACATCTGTAGCTGTGCTTTGTTGATC	240
Qy	241	TTGTTGACTATGAGGAAGTTCTAGGAAACCTAGGAATCTATGATGCTGATGGATGGAG	300
Db	241	TTGTTGACTATGAGGAAGTTCTAGGAAACCTAGGAATCTATGATGCTGATGGATGGAG	300
Qy	301	ATTTTGTGATGGATGATGCCAAAGTTTATTTAGGACTTAAAGAGAGATCTACTTCAGAGC	360
Db	301	ATTTTGTGATGGATGATGCCAAAGTTTATTTAGGACTTAAAGAGAGATCTACTTCAGAGC	360
Qy	361	CAGCAGTCCCAGGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGTTCCCTGG	420
Db	361	CAGCAGTCCCAGGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGTTCCCTGG	420
Qy	421	AGGCAGAACCCAGAAATATCGAAGATGAGCAAAAGAACTTGCACAAAGAGATGGACCCACAG	480
Db	421	AGGCAGAACCCAGAAATATCGAAGATGAGCAAAAGAACTTGCACAAAGAGATGGACCCACAG	480
Qy	481	AAATGGTACGCGAGACATGCTGAGGAGAGAGCTTGCACAAAGAGATGGACCCACAG	540
Db	481	AAATGGTACGCGAGACATGCTGAGGAGAGAGCTTGCACAAAGAGATGGACCCACAG	540
Qy	541	GAGAACCAACAAAGAGGATGATGATTTCTTATGGGACTGATGTAGATGATGATTTG	600
Db	541	GAGAACCAACAAAGAGGATGATGATTTCTTATGGGACTGATGTAGATGATGATTTG	600
Qy	601	AGACCCCTGGAAACCTGAAGTATCTCATGAGAAACCGAGCATAGTTACCACTGGGAGAGA	660

Db	601	AGACCCCTGGAAACCTGAAGTATCTCATGAGAAACCGAGCATAGTTACCACTGGGAGAGA	660
Qy	661	CAGTTTCAAGAGACTCTAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG	720
Db	661	CAGTTTCAAGAGACTCTAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG	720
Qy	721	ATTCCAGTGAACCAAGTAGTAGAAGATTGCAACCATATGACAGATGATGTAAACAT	780
Db	721	ATTCCAGTGAACCAAGTAGTAGAAGATTGCAACCATATGACAGATGATGTAAACAT	780
Qy	781	ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAATGAGGATAGAAATCA	840
Db	781	ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAATGAGGATAGAAATCA	840
Qy	841	CAGAAGTAACTGCTCCCTGAGGATAATCTCTAGAAATTCACAGGTAATTTGTGAAG	900
Db	841	CAGAAGTAACTGCTCCCTGAGGATAATCTCTAGAAATTCACAGGTAATTTGTGAAG	900
Qy	901	AAGTAAGCAATTTTCTCTGTGGAAGAAACAGCAGAAAGTACACAGAAACAAATAGAAAA	960
Db	901	AAGTAAGCAATTTTCTCTGTGGAAGAAACAGCAGAAAGTACACAGAAACAAATAGAAAA	960
Qy	961	CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAGAGAGCCTTAAATAAAT	1020
Db	961	CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAGAGAGCCTTAAATAAAT	1020
Qy	1021	TTGATAAGACTATTAAGCTGAACTTGATGCTCGAGAAACCTCCGTAAAGGGGAAAA	1080
Db	1021	TTGATAAGACTATTAAGCTGAACTTGATGCTCGAGAAACCTCCGTAAAGGGGAAAA	1080
Qy	1081	TTGAGGAAGCAGTGAATGCAATTAAGAACTAGTAGCAAAATACCTCAGAGTCCAGAG	1140
Db	1081	TTGAGGAAGCAGTGAATGCAATTAAGAACTAGTAGCAAAATACCTCAGAGTCCAGAG	1140
Qy	1141	CAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGTAATCAGG	1200
Db	1141	CAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGTAATCAGG	1200
Qy	1201	TGCTAGTGGAGGCATCGAGACCTCAAGAGGTGGCCAGCCTACCTGATGTCCTCCAG	1260
Db	1201	TGCTAGTGGAGGCATCGAGACCTCAAGAGGTGGCCAGCCTACCTGATGTCCTCCAG	1260
Qy	1261	ACCTGCTGAAGCTGAGTTGAAAGCTGCTCAGACAGGCAAACTTCTAGGTCATATGA	1320
Db	1261	ACCTGCTGAAGCTGAGTTGAAAGCTGCTCAGACAGGCAAACTTCTAGGTCATATGA	1320
Qy	1321	GAGTTCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCCTTAA	1380
Db	1321	GAGTTCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCCTTAA	1380
Qy	1381	AAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAAAGAAAGTTT	1440
Db	1381	AAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAAAGAAAGTTT	1440
Qy	1441	ATGAGAGGCTGCTGAGTGTGACACCTTAATGATGGCTTGTGCTTAAAGTCCATTTGCTTCA	1500
Db	1441	ATGAGAGGCTGCTGAGTGTGACACCTTAATGATGGCTTGTGCTTAAAGTCCATTTGCTTCA	1500
Qy	1501	TCCTGAGGACACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAGGATAGAAAT	1560
Db	1501	TCCTGAGGACACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAGGATAGAAAT	1560
Qy	1561	CCGAGATCTCTGGCACTGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGAGA	1620
Db	1561	CCGAGATCTCTGGCACTGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGAGA	1620
Qy	1621	GGGTTGGGAAACAAAGAGGATTAAGTGTGATGCTTGGGCAACAGAGGACACTTTG	1680
Db	1621	GGGTTGGGAAACAAAGAGGATTAAGTGTGATGCTTGGGCAACAGAGGACACTTTG	1680
Qy	1681	CATCTGTCTGGCAACGCTCACTCTACAAATGTAATGAGCATGAAAGCAGCAGCTTGGTGA	1740
Db	1681	CATCTGTCTGGCAACGCTCACTCTACAAATGTAATGAGCATGAAAGCAGCAGCTTGGTGA	1740

1681 CATCTGTGCGCAACGCTCACTCAATGATGTAATGGAATGAAAGCACAGCCTTGGTGA 1740  
1741 CCCCAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGAGTTAA 1800  
1741 CCCCAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGAGTTAA 1800  
1801 TCCGAGATGAAGGCTTGCAGTGTGATGAATGAAGCCAAAGGTCCTTCCCTGCTGAGGATG 1860  
1801 TCCGAGATGAAGGCTTGCAGTGTGATGAATGAAGCCAAAGGTCCTTCCCTGCTGAGGATG 1860  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGCGCAGCAAGGAAGAA 1920  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGCGCAGCAAGGAAGAA 1920  
1921 ATGAAATGCTGCAAGAGGCTCCTTAAACCTGTACTAGAAAAAGTTCCCGGAGA 1980  
1921 ATGAAATGCTGCAAGAGGCTCCTTAAACCTGTACTAGAAAAAGTTCCCGGAGA 1980  
1981 CAACAGATGCAAGAGGACAGATCAATATTCATCATGCAACCCGGGACTCAGGTG 2040  
1981 CAACAGATGCAAGAGGACAGATCAATATTCATCATGCAACCCGGGACTCAGGTG 2040  
2041 GCGCGCACAGAGGCCCCAACAATGAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
2041 GCGCGCACAGAGGCCCCAACAATGAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
2101 AGGAAGGCTGCAAGATTCGATGTGCAACGAGACCCAGGACCTGGGAGGAAAGCAAGTGC 2160  
2101 AGGAAGGCTGCAAGATTCGATGTGCAACGAGACCCAGGACCTGGGAGGAAAGCAAGTGC 2160  
2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGCGAGGATGCTCATCTTCCGCTGA 2220  
2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGCGAGGATGCTCATCTTCCGCTGA 2220  
2221 TATTTCATGCTGATGTGGGATCGGATCGGAATGCAACACAGCAGAGCAGGCTTCGAG 2280  
2221 TATTTCATGCTGATGTGGGATCGGATCGGAATGCAACACAGCAGAGCAGGCTTCGAG 2280  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 2

US-09-903-199-3  
; Sequence 3, Application US/09903199  
; Patent No. 6797496  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV4  
; CURRENT APPLICATION NUMBER: US/09/903,199  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-199-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 CGGACCGTGCATGCGCCAGCGTGAAGATGCCAAGAGCGGCGCACAGCAGCGAGG 60  
1 CGGACCGTGCATGCGCCAGCGTGAAGATGCCAAGAGCGGCGCACAGCAGCGAGG 60

61 GCTCCGACAGCGGTAGCACAGGTGCGGCGACGACGAGCCCGGGGCGCGAGAGACAA 120  
61 GCTCCGACAGCGGTAGCACAGGTGCGGCGACGACGAGCCCGGGGCGCGAGAGACAA 120  
121 AGCATGAGAGACACAAAGATGGAGAAAGCGGAGCTCTCGGGAACCTTATCTTCAAGT 180  
121 AGCATGAGAGACACAAAGATGGAGAAAGCGGAGCTCTCGGGAACCTTATCTTCAAGT 180  
181 GGTATTATGGTGAATGTCATTGCTGGGCGCTGCGACATCTGTAGCTGTCTGTTGGTTGATC 240  
181 GGTATTATGGTGAATGTCATTGCTGGGCGCTGCGACATCTGTAGCTGTCTGTTGGTTGATC 240  
241 TTGTTGACTATGAGGAAGTTCTAGGAAACCTAGGAATCTATGATGCTGTGATGGTGAAG 300  
241 TTGTTGACTATGAGGAAGTTCTAGGAAACCTAGGAATCTATGATGCTGTGATGGTGAAG 300  
301 ATTTTGTATGTTGATGATGCTCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTCAAGC 360  
301 ATTTTGTATGTTGATGATGCTCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTCAAGC 360  
361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACATGAGCCCGGAGGAGCAGGTTCTCTGTGG 420  
361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACATGAGCCCGGAGGAGCAGGTTCTCTGTGG 420  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
481 AAATGTTGATGATGATGCTCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTCAAGC 540  
481 AAATGTTGATGATGATGCTCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTCAAGC 540  
541 GAGAACCAACCAAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATG 600  
541 GAGAACCAACCAAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATG 600  
601 AGACCTTGGAACTTCAAGTATCTCATGAAGAAACCGAGCATAGTTACCGTGGAAAGAGA 660  
601 AGACCTTGGAACTTCAAGTATCTCATGAAGAAACCGAGCATAGTTACCGTGGAAAGAGA 660  
661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGCTGTGAGCAGGAAATCCAG 720  
661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGCTGTGAGCAGGAAATCCAG 720  
721 ATTCAGTGAACCAAGTATGATGATGAAGATGCAACCATGATACAGATGATGATGATG 780  
721 ATTCAGTGAACCAAGTATGATGATGAAGATGCAACCATGATACAGATGATGATGATG 780  
781 ACCAAGTCTATGAGGAAACCAAGTATGATGATGATGATGATGATGATGATGATGATG 840  
781 ACCAAGTCTATGAGGAAACCAAGTATGATGATGATGATGATGATGATGATGATGATG 840  
841 CAGAAGTAACTGCTCCCTGAGGATAATCTCTGTAGAAGATTCACAGGTAATTTGTAGAAG 900  
841 CAGAAGTAACTGCTCCCTGAGGATAATCTCTGTAGAAGATTCACAGGTAATTTGTAGAAG 900  
901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGGATACCAAGAAACAAATAGAAAAA 960  
901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGGATACCAAGAAACAAATAGAAAAA 960  
961 CAGATGATCCAGAACCAAAAGTTTAAAGAAAGAGGCTTAAACTTTTAAATTAAT 1020  
961 CAGATGATCCAGAACCAAAAGTTTAAAGAAAGAGGCTTAAACTTTTAAATTAAT 1020  
1021 TTGATAAGACTATTTAAAGCTGAATCTGTGTCAGAAAAAATCCCTGAGGTTCCACGAG 1080  
1021 TTGATAAGACTATTTAAAGCTGAATCTGTGTCAGAAAAAATCCCTGAGGTTCCACGAG 1080  
1081 TTGAGGAAGCAGTGAATGATGATTTAAAGAACTAGTACGCAATATCCCTCAGAGTCCACGAG 1140  
1081 TTGAGGAAGCAGTGAATGATGATTTAAAGAACTAGTACGCAATATCCCTCAGAGTCCACGAG 1140

```
QY 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGTGAGAGAGAGAAATGAGG 1200
Db 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGTGAGAGAGAGAAATGAGG 1200
QY 1201 TGCTACCTGGAGCCATCGAGACCTTACCAAGAGTGGCCAGCCTACCTGATGTCCTTCAG 1260
Db 1201 TGCTACCTGGAGCCATCGAGACCTTACCAAGAGTGGCCAGCCTACCTGATGTCCTTCAG 1260
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGGCTCGCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGGCTCGCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320
QY 1321 GAGGTTCCCTGCTTACCCCTGCAGAGATTAGTTCAACTATTTCCTCAATGATACCTTCCTTAA 1380
Db 1321 GAGGTTCCCTGCTTACCCCTGCAGAGATTAGTTCAACTATTTCCTCAATGATACCTTCCTTAA 1380
QY 1381 AAATGACCTTGGCGTGGGATACCTCTTGATAGAGATATGCAATGCCAAGAAAGTTT 1440
Db 1381 AAATGACCTTGGCGTGGGATACCTCTTGATAGAGATATGCAATGCCAAGAAAGTTT 1440
QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500
Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500
QY 1501 TCCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCCATATTAAAGAAAGGAATAGAT 1560
Db 1501 TCCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCCATATTAAAGAAAGGAATAGAT 1560
QY 1561 CCGGAGATCTGGCACTGATGATGGGAGATTTATTTCCACCTGGGGGATGCCATGCGAGA 1620
Db 1561 CCGGAGATCTGGCACTGATGATGGGAGATTTATTTCCACCTGGGGGATGCCATGCGAGA 1620
QY 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCACAAAGAGAGACATTTTG 1680
Db 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCACAAAGAGAGACATTTTG 1680
QY 1681 CATCTGCTGCGAACGCTCTACTTACAAATGTGAATGAGCTGAAAGCAACAGCTTTGGTGA 1740
Db 1681 CATCTGCTGCGAACGCTCTACTTACAAATGTGAATGAGCTGAAAGCAACAGCTTTGGTGA 1740
QY 1741 CCCCAGAGAAACGGGCTACACAGATTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800
Db 1741 CCCCAGAGAAACGGGCTACACAGATTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800
QY 1801 TCCGAGATGAAGGCTTTGCACTGATGATGAATGAAGCAAGGCTCTCTCTGCTGAGGATG 1860
Db 1801 TCCGAGATGAAGGCTTTGCACTGATGATGAATGAAGCAAGGCTCTCTCTGCTGAGGATG 1860
QY 1861 AAAACCTGAGGGAAGGAGCTGGAGCCAGTTTCAAGCTGTGGCAGCAAGAAAGAA 1920
Db 1861 AAAACCTGAGGGAAGGAGCTGGAGCCAGTTTCAAGCTGTGGCAGCAAGAAAGAA 1920
QY 1921 ATGAAGATGCTGCAAGAGGCTCTTAAACCTGCTACTAGTAAAGTTTCCCCGAGA 1980
Db 1921 ATGAAGATGCTGCAAGAGGCTCTTAAACCTGCTACTAGTAAAGTTTCCCCGAGA 1980
QY 1981 CAACAGGATGAGAAAGGAGCAGATCAATATTCATATCCATGTCACCCCGGACTCACGTTG 2040
Db 1981 CAACAGGATGAGAAAGGAGCAGATCAATATTCATATCCATGTCACCCCGGACTCACGTTG 2040
QY 2041 GGCGCGACACAGGCGCCCAAACTGCGAGGCTCCGAATGTCACCTGGGCTTGTGATCCCA 2100
Db 2041 GGCGCGACACAGGCGCCCAAACTGCGAGGCTCCGAATGTCACCTGGGCTTGTGATCCCA 2100
QY 2101 AGGAAGCTGCAAGATTCGATGTCGCAACGAGACCGAGCCTGGAGGAGGCAAGGTGC 2160
Db 2101 AGGAAGCTGCAAGATTCGATGTCGCAACGAGACCGAGCCTGGAGGAGGCAAGGTGC 2160
QY 2161 TCATCTTTTGTGACTCTCTTTGAGCAGCAGGATGATGGCAGGATGCTCATCTTTCCGGCTGA 2220
Db 2161 TCATCTTTTGTGACTCTCTTTGAGCAGCAGGATGATGGCAGGATGCTCATCTTTCCGGCTGA 2220
QY 2221 TATTTCATGCTGGATGTGTGGCATCCGGNACTGACACACACAGCAGAGACGCGCTTCCAG 2280
```

```
Db 2221 TATTTCATGCTGGATGTGTGGCATCCGGAACCTGACACACAGCAGACGAGCCTTCCAG 2280
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGAAAACTCTGGAGAGA 2324
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGAAAACTCTGGAGAGA 2324

RESULT 3
US-09-903-216-3
; Sequence 3, Application US/09903216
; Patent No. 6812206
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M.
; APPLICANT: Ince, Nedim
; APPLICANT: Carlson, Rolf I.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 DIV2
; CURRENT APPLICATION NUMBER: US/09/903,216
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2324
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-903-216-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATGGCCCGCCAGCTTAAGATGCCAAGAGCAGCGGCCCAACAGCAGCAGCAGCG 60
Db 1 CGGACCGTGCATGGCCCGCCAGCTTAAGATGCCAAGAGCAGCGGCCCAACAGCAGCAGCAGCG 60
QY 61 GCTCCGCGCAGCGGTAGCACAGTGGCGGAGCAGCAGCCCGCGGCCCGCGAGAGAGACAA 120
Db 61 GCTCCGCGCAGCGGTAGCACAGTGGCGGAGCAGCAGCCCGCGGCCCGCGAGAGAGACAA 120
QY 121 AGCATGGAGAGACAAAGAATGGAGAAAGGCGGACTCTCGGGAACCTTCACTTTCACGT 180
Db 121 AGCATGGAGAGACAAAGAATGGAGAAAGGCGGACTCTCGGGAACCTTCACTTTCACGT 180
QY 181 GGTTCATGCTGATGCTGCTGGCGCTCTGGACATCTGATCTGCTGTTGGTTTCATC 240
Db 181 GGTTCATGCTGATGCTGCTGGCGCTCTGGACATCTGATCTGCTGTTGGTTTCATC 240
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGGTGATGGAG 300
Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGGTGATGGAG 300
QY 301 ATTTTGTGATGATGTCGCAAGTTTTATAGGACTTAAAGAGAGATCTACTTTCAGAGC 360
Db 301 ATTTTGTGATGATGTCGCAAGTTTTATAGGACTTAAAGAGAGATCTACTTTCAGAGC 360
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGTTCCTGTGG 420
Db 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGTTCCTGTGG 420
QY 421 AGGCAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480
Db 421 AGGCAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480
QY 481 AAATGGTACACGAGAAACATGTTGAGGAGAGAGACTTGGCAACAAGAGATGGACCCACAG 540
Db 481 AAATGGTACACGAGAAACATGTTGAGGAGAGAGACTTGGCAACAAGAGATGGACCCACAG 540
QY 541 GAGAACCAACACAGAGGATGATGATGCTTCTTATGGCGACTGATGATAGATGATGATTTG 600
```

Db 541 GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTGG 600  
Qy 601 AGACCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACCATGTTGAAGAGA 660  
Db 601 AGACCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACCATGTTGAAGAGA 660  
Qy 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACCAAGTATGAGAGTGAAGAGTTTGACCATGATACAGATGATGTAACAT 780  
Db 721 ATTCCAGTGAACCAAGTATGAGAGTGAAGAGTTTGACCATGATACAGATGATGTAACAT 780  
Qy 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAATGAAAGGATGAATCA 840  
Db 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAATGAAAGGATGAATCA 840  
Qy 841 CAGAAGTAACTGCTCCCTCAGGATAATCCTGTGAGAGATTTCAAGGTAAATTTGAGAG 900  
Db 841 CAGAAGTAACTGCTCCCTCAGGATAATCCTGTGAGAGATTTCAAGGTAAATTTGAGAG 900  
Qy 901 AAGTAAAGCATTTTCTGTGGAAGAACAGCAGGAGAGTACCAACAGAAACAAATAGAAAAA 960  
Db 901 AAGTAAAGCATTTTCTGTGGAAGAACAGCAGGAGAGTACCAACAGAAACAAATAGAAAAA 960  
Qy 961 CAGATGATCCAGAAACAAAGCAAAAGTTTAAAGAAAGAACCTTAAATTAATTAAT 1020  
Db 961 CAGATGATCCAGAAACAAAGCAAAAGTTTAAAGAAAGAACCTTAAATTAATTAAT 1020  
Qy 1021 TTGATAAGACTATTAAGCTGAACCTGATGCTGCGAAGAAACCTCGTAAAGGGGAAAAA 1080  
Db 1021 TTGATAAGACTATTAAGCTGAACCTGATGCTGCGAAGAAACCTCGTAAAGGGGAAAAA 1080  
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
Qy 1141 CAAGATATGGGAAGCCGACGTGTGAGGATGATTTGGCTGAGAAGAGAGAAAGTAATGAGG 1200  
Db 1141 CAAGATATGGGAAGCCGACGTGTGAGGATGATTTGGCTGAGAAGAGAGAAAGTAATGAGG 1200  
Qy 1201 TGCTACGTGAGGACCATCGAGACCTACCAAGAGGTGGCAGCCTACCTGATGCTCCGCGAG 1260  
Db 1201 TGCTACGTGAGGACCATCGAGACCTACCAAGAGGTGGCAGCCTACCTGATGCTCCGCGAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCCGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCCGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
Qy 1321 GAGGTTCCTCTTACCCTCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTCTTAA 1380  
Db 1321 GAGGTTCCTCTTACCCTCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTCTTAA 1380  
Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCAAGAAAGATTT 1440  
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCAAGAAAGATTT 1440  
Qy 1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTGTCTAAAGTCCATATGCGTTCA 1500  
Db 1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTGTCTAAAGTCCATATGCGTTCA 1500  
Qy 1501 TCCTGAAGGCAAGAAATTTGCTGAGAGCATCCCATATTTAAAGAGGAAGTAAGAT 1560  
Db 1501 TCCTGAAGGCAAGAAATTTGCTGAGAGCATCCCATATTTAAAGAGGAAGTAAGAT 1560  
Qy 1561 CCGGAGATCTCGGACCTGATGATGGAGATTTTATTTCCACTCGGGGATGCCATCGACA 1620  
Db 1561 CCGGAGATCTCGGACCTGATGATGGAGATTTTATTTCCACTCGGGGATGCCATCGACA 1620  
Qy 1621 GGGTTGGGAACAAGAGGCATATAAGTGTATGAGCTTGGGCAACAAGAGGACATTTG 1680  
Db 1621 GGGTTGGGAACAAGAGGCATATAAGTGTATGAGCTTGGGCAACAAGAGGACATTTG 1680

Qy 1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGGAATGAAGCAGCAGCTTGGTGA 1740  
Db 1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGGAATGAAGCAGCAGCTTGGTGA 1740  
Qy 1741 CCCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACTGGAGTTAA 1800  
Db 1741 CCCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACTGGAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTGCGAGTATGATGAATAAAGCCAAAGGTTCTTCTCTGCTCAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTGCGAGTATGATGAATAAAGCCAAAGGTTCTTCTCTGCTCAGGATG 1860  
Qy 1861 AAAACCTCAGGGAAGAAAGGGGACTGGAGCCAGTTCAAGCTGTGGCAGCAAGGAAGAA 1920  
Db 1861 AAAACCTCAGGGAAGAAAGGGGACTGGAGCCAGTTCAAGCTGTGGCAGCAAGGAAGAA 1920  
Qy 1921 ATGAAAATGCTCTCAAAAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGAGA 1980  
Db 1921 ATGAAAATGCTCTCAAAAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGAGA 1980  
Qy 1981 CAACAGGATGCAAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGAGTCACTG 2040  
Db 1981 CAACAGGATGCAAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGAGTCACTG 2040  
Qy 2041 GSCCCGACACAGGSCCCACAAACTGCAGGCTCCGAATGCACCTGGCTTGGTGAATCCCA 2100  
Db 2041 GSCCCGACACAGGSCCCACAAACTGCAGGCTCCGAATGCACCTGGCTTGGTGAATCCCA 2100  
Qy 2101 AGGAAGGCTCAAGATTCGATGTGCCAACAGACAGACAGCCTGGGAGAGGCAAGTGC 2160  
Db 2101 AGGAAGGCTCAAGATTCGATGTGCCAACAGACAGACAGCCTGGGAGAGGCAAGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTTGACACGAGGTATGCGAGATGCTCATCTTTCCGGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTTGACACGAGGTATGCGAGATGCTCATCTTTCCGGCTGA 2220  
Qy 2221 TATTCATCTGATGTGTGGCATCCGAACTGCACACAGCAGACAGCAGCCTTCCAG 2280  
Db 2221 TATTCATCTGATGTGTGGCATCCGAACTGCACACAGCAGACAGCAGCCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATCAAGCTTGGGAACCTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATCAAGCTTGGGAACCTCTGGAGAGA 2324

## RESULT 4

US-09-903-063-3  
; Sequence 3, Application US/09903063  
; Patent No. 6815415  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV3  
; CURRENT APPLICATION NUMBER: US/09/903,063  
; PRIOR FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-063-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Db 2161 TCATCTTTGATGACTCTCTTTGAGCAGAGGATATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
Qy 2221 TATTTCATCGTGGATGTGTGGCATCCGGAATGACACACACAGCAGACGCGCTTCCAG 2280  
Db 2221 TATTTCATCGTGGATGTGTGGCATCCGGAATGACACACACAGCAGACGCGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGAACTCTGGAGAGA 2324

## RESULT 5

US-09-859-604-3  
; Sequence 3, Application US/09859604  
; Patent No. 6835370  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M  
; APPLICANT: Deutch, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/436,184  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-859-604-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCTGCAATGGCCCGCAGCTAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCTGCAATGGCCCGCAGCTAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGCGAGCGGTAGCAGCAGAGTCGCGGCGCAGCAGCAGCCCGGGGCCCCGAGAGAGACAA 120  
Db 61 GCTCCGCGAGCGGTAGCAGCAGAGTCGCGGCGCAGCAGCAGCCCGGGGCCCCGAGAGAGACAA 120  
Qy 121 AGCATGGAGGACACAGAATGGGAGGAAAGGCGGACTCTCGGGAACCTTCAATCTTCACGT 180  
Db 121 AGCATGGAGGACACAGAATGGGAGGAAAGGCGGACTCTCGGGAACCTTCAATCTTCACGT 180  
Qy 181 GGTATTGATGATTTGATGCTGGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC 240  
Db 181 GGTATTGATGATTTGATGCTGGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC 240  
Qy 241 TTGTTGACTATGAGGAGTCTAGGAAACTAGGAATCTATGATGCTGTGATGCTGATGAG 300  
Db 241 TTGTTGACTATGAGGAGTCTAGGAAACTAGGAATCTATGATGCTGTGATGCTGATGAG 300  
Qy 301 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 301 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG 420  
Qy 421 AGGCGAGACCCGAGATATCGAAGATGAGCAAGAAAGAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGCGAGACCCGAGATATCGAAGATGAGCAAGAAAGAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGCAAGAACATGTTGAGGGAGAGAGATTCGCAACAGAAAGATGAGCCACAG 540  
Db 481 AAATGGTACACGCAAGAACATGTTGAGGGAGAGAGATTCGCAACAGAAAGATGAGCCACAG 540

Qy 541 GAGAACCAACCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Qy 601 AGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTATACCATGCGAAGAGA 660  
Db 601 AGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTATACCATGCGAAGAGA 660  
Qy 661 CAGTTTCAAGAGCTGTAAATCAGGATATGGAAGAGATGATGCTGTGACGAGAAATCCAG 720  
Db 661 CAGTTTCAAGAGCTGTAAATCAGGATATGGAAGAGATGATGCTGTGACGAGAAATCCAG 720  
Qy 721 ATTCAGTGAACCAAGTAGTAGAAGATGAAAGATTTGCCACCATGATACAGATGATGATACAT 780  
Db 721 ATTCAGTGAACCAAGTAGTAGAAGATGAAAGATTTGCCACCATGATACAGATGATGATACAT 780  
Qy 781 ACCAAGTCTATGAGGAAACCAAGCATATATGAACTCTAGAAAATGAAAGGATGAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAAACCAAGCATATATGAACTCTAGAAAATGAAAGGATGAGAAATCA 840  
Qy 841 CAGAAGTAACTGCTCCCTCCCTGAGGATTAATCTCTGTAGAAGATTCACAGGTAAATTTGTAAG 900  
Db 841 CAGAAGTAACTGCTCCCTCCCTGAGGATTAATCTCTGTAGAAGATTCACAGGTAAATTTGTAAG 900  
Qy 901 AAGTAAGCATTTTCTGTGGAAGAACAGCATGAGTAAAGATTAAGAAAAGAGCCCTAAATTAAT 1020  
Db 901 AAGTAAGCATTTTCTGTGGAAGAACAGCATGAGTAAAGATTAAGAAAAGAGCCCTAAATTAAT 1020  
Qy 1021 TTGTAAGACTTATTAAGCTGAACTTGAATGCTGTGCAAGAAACTAGTACGCAAAATACCTC 1080  
Db 1021 TTGTAAGACTTATTAAGCTGAACTTGAATGCTGTGCAAGAAACTAGTACGCAAAATACCTC 1080  
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
Qy 1141 CAAGATATGGAAGGCGAGTGTGAGATGATTTGGCTGCAAGAGAGAGATTAATGAGG 1200  
Db 1141 CAAGATATGGAAGGCGAGTGTGAGATGATTTGGCTGCAAGAGAGAGATTAATGAGG 1200  
Qy 1201 TGCTACGTGAGGCGCATCGAGACCTACCAAGAGGTTGGCCAGCTACCTGATGTCCCTGCAG 1260  
Db 1201 TGCTACGTGAGGCGCATCGAGACCTACCAAGAGGTTGGCCAGCTACCTGATGTCCCTGCAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGTGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGTGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
Qy 1321 GAGGTTCCCTGCTTACCTGCGAGATTAAGTAACTATTTCCCAATGATACTTCCCTTAA 1380  
Db 1321 GAGGTTCCCTGCTTACCTGCGAGATTAAGTAACTATTTCCCAATGATACTTCCCTTAA 1380  
Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATTAAGCAATGCAAGAAAGATTT 1440  
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATTAAGCAATGCAAGAAAGATTT 1440  
Qy 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGCTTTGCTTAAAGTCCATTTATGCTTCA 1500  
Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGCTTTGCTTAAAGTCCATTTATGCTTCA 1500  
Qy 1501 TCCTGAAGGCGACAGAAACAAATTTGCTGAGAGCATCCCAATTTAAAGGAAAGGAATAGAAT 1560  
Db 1501 TCCTGAAGGCGACAGAAACAAATTTGCTGAGAGCATCCCAATTTAAAGGAAAGGAATAGAAT 1560  
Qy 1561 CCGGAGATCCCTGGCAGTGTGAGGAGATTTTATTTTCCACCTGGGGATGCCATGCGAGA 1620  
Db 1561 CCGGAGATCCCTGGCAGTGTGAGGAGATTTTATTTTTCATCTGGGGATGCCATGCGAGA 1620

1621 GGGTTGGGAAACAAGAGGAGCATATAGTGGTATAGCTTTGGGCAACAAGAGGAGCACTTTTG 1680  
Db GGGTTGGGAAACAAGAGGAGCATATAGTGGTATAGCTTTGGGCAACAAGAGGAGCACTTTTG 1680  
1681 CATCTGCTGCAACGCTCACTACATGATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1740  
Db CATCTGCTGCAACGCTCACTACATGATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1740  
1741 CCCCCAAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db CCCCCAAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1801 TCCGAGATGAAGGCTTGCAGTATGGAATGAAGCAAGGCTCTTTCTGCTGAGGATG 1860  
Db TCCGAGATGAAGGCTTGCAGTATGGAATGAAGCAAGGCTCTTTCTGCTGAGGATG 1860  
1861 AAAACCTGAGGAAAGGAGGACTGGAGCCAGTTTCAAGCTGTGCGAGCAAGGAAGAA 1920  
Db AAAACCTGAGGAAAGGAGGACTGGAGCCAGTTTCAAGCTGTGCGAGCAAGGAAGAA 1920  
1921 ATGAAATGCTGCAAGAGGAGCTCTAAACCTGTACCTTACTAGAAAGTTTCCCGAGA 1980  
Db ATGAAATGCTGCAAGAGGAGCTCTAAACCTGTACCTTACTAGAAAGTTTCCCGAGA 1980  
1981 CAACAGGATGAGAGGAGGAGAGATCAATATTCATCATGACACCCGGGAGTCACTGT 2040  
Db CAACAGGATGAGAGGAGGAGAGATCAATATTCATCATGACACCCGGGAGTCACTGT 2040  
2041 GGCCGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
Db GGCCGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
2101 AGAAGGCTGCAAGATTCGATGTGCAACAGAGACAGGAGGAGGAGGAGGAGGAGGAGG 2160  
Db AGAAGGCTGCAAGATTCGATGTGCAACAGAGACAGGAGGAGGAGGAGGAGGAGGAGG 2160  
2161 TCATCTTTGATGACTCTTTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220  
Db TCATCTTTGATGACTCTTTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220  
2221 TATTTCATGCTGATGTGGGATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280  
Db TATTTCATGCTGATGTGGGATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280  
2281 CAATTAGCATGAATTCATGAAGCTTGGGAACTCTGGAGAG 2324  
Db CAATTAGCATGAATTCATGAAGCTTGGGAACTCTGGAGAG 2324

RESULT 6  
US-09-040-485-1  
Sequence 1, Application US/09040485  
Patent No. 6166176  
GENERAL INFORMATION:  
APPLICANT: Radosevich, James A.  
TITLE OF INVENTION: A GENE ENCODING A NOVEL MARKER FOR  
TITLE OF INVENTION: CANCER  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60611-5599  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/040,485

FILING DATE: 17-MAR-1998  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Martin, Alice O.  
REGISTRATION NUMBER: 35,601  
REFERENCE/DOCKET NUMBER: 8998/3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-321-4200  
TELEFAX: 312-321-4299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2442 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 70..834  
US-09-040-485-1  
Query Match 33.8%; Score 785; DB 3; Length 2442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 162 GGAACCTTCATCTTCACGTGGTGTATGCTGATGCTGCTGGCGCTGCGACATCTGTA 221  
Db 46 GGAACCTTCATCTTCACGTGGTGTATGCTGATGCTGCTGGCGCTGCGACATCTGTA 105  
Qy 222 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 281  
Db 106 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 165  
Qy 282 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 341  
Db 166 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 225  
Qy 342 GAGAGATCTACTTCAGAGCCAGAGTCCCGCCAGAGAGGCTGAGCCACACACTTGAGGCC 401  
Db 226 GAGAGATCTACTTCAGAGCCAGAGTCCCGCCAGAGAGGCTGAGCCACACACTTGAGGCC 285  
Qy 402 GAGAGAGAGGCTTCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 461  
Db 286 GAGAGAGAGGCTTCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 345  
Qy 462 ATTCAGTCCCTTCCTCCATGAAATGGTACACGCAACATGTTGAGGAGGAGGAGGAGGAG 521  
Db 346 ATTCAGTCCCTTCCTCCATGAAATGGTACACGCAACATGTTGAGGAGGAGGAGGAGGAG 405  
Qy 522 CAAGAAGATGGACCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 581  
Db 406 CAAGAAGATGGACCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 465  
Qy 582 GAT 641  
Db 466 GAT 525  
Qy 642 AGTTACCACTGGAG 701  
Db 526 AGTTACCACTGGAG 585  
Qy 702 TCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 761  
Db 586 TCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 645  
Qy 762 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 821  
Db 646 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 705  
Qy 822 AATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 881  
Db 706 AATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 765



**Qy**

882 TCACAGGTAATTGTAGAAGATGAAGCATTTTCTCGTGGAAGAACAGCAGCAAGATACC A 941  
|||||  
|||

**Dd**

766 TCACAGGTAATTGTAGAAGATGAAGCATTTTCTCGTGGAAGAACAGCAGCAAGATACC A 825  
|||||  
|||

Qy 942 CCAGA 946  
Db 826 CCAGA 830

## RESULT 7

US-09-702-705-1791  
; Sequence 1791, Application US/09702705

: Patent No. 6504010

; APPLICANT: Banqur, Chaitanya S.

/ APPLICANT: Lodes, Michael A.  
/  
/ APPLICANT: Fanger, Gary  
/  
/ APPLICANT: Vedvick, Tom  
/  
/ APPLICANT: Carter, Parrick  
/  
/ APPLICANT: Retter, Marc  
/  
/ APPLICANT: Mannion, Jane  
/  
/ APPLICANT: Fan, Liguon

APPLICANT: FAIR  
: TITLE OF INVENTION

```

1  TITLE OF INVENTION: COMPOSITIONS AND METHODS
2  FOR THE THERAPY AND
3  TREATMENT OF
4  TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
5
6  FILE REFERENCE: 210121.478C14
7
8  CURRENT APPLICATION NUMBER: US/09/702,705
9
10 CURRENT FILING DATE: 2000-10-30
11
12 NUMBER OF SEQ ID NOS: 1833
13
14 SOFTWARE: FastSeq for Windows Version 3.0
15
16 SEQ ID NO 1791
17
18 LENGTH: 2442
19
20 TYPE: DNA
21
22 ORGANISM: Homo sapiens
23
24 US-09-702-705-1791

```

Db	526		AGTTACACAGTGAAGAGACAGTTTCAAGAGCTGTAATCAGAGTATGGAAGAGTATG	585
Qy	702		TCTGAGCAGAGAAATCCAGATTCCAGATGAACCACTAGTAGAAGATTGAAAGATTGCACCAT	761
Db	586		TCTGAGCAGAGAAATCCAGATTCCAGTGAACCACTAGTAGAAGATTGAAAGATTGCACCAT	645
Qy	762		GATACACAGATGTTATACATACCAGTCTATGCGACACAGCAGTATATGACCTCTTAGAA	821
Db	646		GATACACAGATGTTATACATACAGTCTATGCGACACAGCAGTATATGACCTCTTAGAA	705
Qy	822		AATGAAGGGATAGAAATCACAGAAGTAACCTG	852
Db	706		AATGAAGGGATAGAAATCACAGAAGTAACCTG	736

## RESULT 8

US-09-736-457-1791  
; Sequence 1791, Application US/09736457

; Patent No. 6509448

; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodges, Michael A.  
 ; APPLICANT: Panger, Gary  
 ; APPLICANT: Vedvick, Tom  
 ; APPLICANT: Carter, Darick  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Fan, Liqun

APPLICANT: Wang. Aijun

```

/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITL OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C15
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736.457
/ CURRENT APPLICATION NUMBER: US/09/736.457
/ CURRENT FILING DATE: 2000-12-13
/ CURRENT FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 1864
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1791
/ SEQ ID NO 1791
/ LENGTH: 2442
/ LENGTH: 2442
/ TYPE: DNA
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ ORGANISM: Homo sapiens
/ US-09-736-457-1791
/ US-09-736-457-1791

```

### Online Match



Db 406 CAAGAAGATGGACCCACAGGAGAACCAACAAGAGGATGATGAGTTCTTATGGCGACT 465  
QY 582 GATGTAGATGATAGATTGAGACCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCAT 641  
Db 466 GATGTAGATGATAGATTGAGACCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCAT 525  
QY 642 AGTTACCAAGTGGAGAGACAGTTTCCAGAGACTGTAAATCAGGATATGGAAGAGATGATG 701  
Db 526 AGTTACCAAGTGGAGAGACAGTTTCCAGAGACTGTAAATCAGGATATGGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAAAATCCAGATTCCAGTGAACAGTAGTAGAAGATGAAGAATTCACCAT 761  
Db 586 TCTGAGCAGGAAAATCCAGATTCCAGTGAACAGTAGTAGAAGATGAAGAATTCACCAT 645  
QY 762 GATACAGATGATGTAACATACCAGTCTATCAGGAAACAGCAGATGATATGAACCTCTAGAA 821  
Db 646 GATACAGATGATGTAACATACCAGTCTATCAGGAAACAGCAGATGATATGAACCTCTAGAA 705  
QY 822 AATGAAGGGATAGAAATCACAGAGTAAGTCTG 852  
Db 706 AATGAAGGGATAGAAATCACAGAGTAAGTCTG 736

## RESULT 9

US-09-671-325-1791

; Sequence 1791, Application US/09671325

; Patent No. 6667154

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedwick, Tom

; APPLICANT: Carter, Darrick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; CURRENT APPLICATION NUMBER: US/09/671,325

; PRIOR FILING DATE: 2000-09-26

; NUMBER OF SEQ ID NOS: 1825

; SOFTWARE: Fast-Seq for Windows Version 3.0

; SEQ ID NO 1791

; LENGTH: 2442

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-671-325-1791

Query Match 27.5%; Score 640; DB 4; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 3e-297;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAACCTTCATCTTCCAGCTGGTTTATGGTGAATTCGATTCGCGGCTGCGACATCTGTA 221  
Db 46 GGAACCTTCATCTTCCAGCTGGTTTATGGTGAATTCGATTCGCGGCTGCGACATCTGTA 105  
QY 222 GCTGCTGTTGGTTTGTGATCTTGTGATGATGAGGAGTTCTAGGAAACTAGGAATCTAT 281  
Db 106 GCTGCTGTTGGTTTGTGATCTTGTGATGATGAGGAGTTCTAGGAAACTAGGAATCTAT 165  
QY 282 GATGCTGATGTTGATGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 166 GATGCTGATGTTGATGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 225  
QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401  
Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 285  
QY 402 GAGAGCAGGTTCTGTGGAGCAGAACCCAGCAATATCGAAGATGAAGCAAGAACAA 461

Db 286 GAGGAGCAGGTTCTGTGGAGGCAAGACCCAGAAATATCGAAGATGAAGCAAAAGAACAA 345  
QY 462 ATTCAAGTCCCTTCTCCATGAAATGCTACACGCAACATGTTGAGGGAGAGACTTGCAG 521  
Db 346 ATTCAAGTCCCTTCTCCATGAAATGCTACACGCAACATGTTGAGGGAGAGACTTGCAG 405  
QY 522 CAAGAAGATGACCCCAAGGAGAACCAACAAGAGGATGATGATGATGATGATGATGATGAT 581  
Db 406 CAAGAAGATGACCCCAAGGAGAACCAACAAGAGGATGATGATGATGATGATGATGATGAT 465  
QY 582 GATGTAGATGATAGATTGAGACCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCAT 641  
Db 466 GATGTAGATGATAGATTGAGACCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCAT 525  
QY 642 AGTTACCAAGTGGAGAGACAGTTTCCAGAGCTGTAAATCAGGATATGGAAGAGATGATG 701  
Db 526 AGTTACCAAGTGGAGAGACAGTTTCCAGAGCTGTAAATCAGGATATGGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAAAATCCAGATTCCAGTGAACAGTAGTAGAAGATGAAGAATTCACCAT 761  
Db 586 TCTGAGCAGGAAAATCCAGATTCCAGTGAACAGTAGTAGAAGATGAAGAATTCACCAT 645  
QY 762 GATACAGATGATGTAACATACCAGTCTATCAGGAAACAGCAGATGATATGAACCTCTAGAA 821  
Db 646 GATACAGATGATGTAACATACCAGTCTATCAGGAAACAGCAGATGATATGAACCTCTAGAA 705  
QY 822 AATGAAGGGATAGAAATCACAGAGTAAGTCTG 852  
Db 706 AATGAAGGGATAGAAATCACAGAGTAAGTCTG 736

## RESULT 10

US-09-659-521-1

; Sequence 1, Application US/09659521

; Patent No. 6727080

; GENERAL INFORMATION:

; APPLICANT: RADOSEVICH, JAMES A.

; TITLE OF INVENTION: GENE ENCODING LABYRINTHIN, A MARKER FOR CANCER

; FILE REFERENCE: 21511/91057

; CURRENT APPLICATION NUMBER: US/09/659,521

; CURRENT FILING DATE: 2000-09-12

; PRIOR APPLICATION NUMBER: PCT/US99/05365

; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: 09/040,485

; PRIOR FILING DATE: 1998-03-17

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2442

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (70)..(834)

US-09-659-521-1

Query Match 27.5%; Score 640; DB 4; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 3e-297;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAACCTTCATCTTCCAGCTGGTTTATGGTGAATTCGATTCGCGGCTGCGACATCTGTA 221  
Db 46 GGAACCTTCATCTTCCAGCTGGTTTATGGTGAATTCGATTCGCGGCTGCGACATCTGTA 105  
QY 222 GCTGCTGTTGGTTTGTGATCTTGTGATGATGAGGAGTTCTAGGAAACTAGGAATCTAT 281  
Db 106 GCTGCTGTTGGTTTGTGATCTTGTGATGATGAGGAGTTCTAGGAAACTAGGAATCTAT 165  
QY 282 GATGCTGATGTTGATGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 166 GATGCTGATGTTGATGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 225  
QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401

Db	226	GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCGAGGAAGGGCTGAGCCACACTGAGCCC	285
Qy	402	GAGGAGCAGGTTCTCTGTGGAGGAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAACAA	461
Db	286	GAGGAGCAGGTTCTCTGTGGAGGAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAACAA	345
Qy	462	ATTCAAGTCCCTTCTCCATGAAATGGTACACGCAGAAACATGTTGAGGAGGAAGACTTTGCCAA	521
Db	346	ATTCAAGTCCCTTCTCCATGAAATGGTACACGCAGAAACATGTTGAGGAGGAAGACTTTGCCAA	405
Qy	522	CAAGAAGATGACCCCAACAGGAGAAACCAACAACAGAGGAGTGAATGAGTTTCTTATGGCGACT	581
Db	406	CAAGAAGATGACCCCAACAGGAGAAACCAACAACAGAGGAGTGAATGAGTTTCTTATGGCGACT	465
Qy	582	GATGTAGATGATAGATTTGAGACCTCTGGAAACCTGAAGTATCTCATGAAGAAACCGAGCAT	641
Db	466	GATGTAGATGATAGATTTGAGACCTCTGGAAACCTGAAGTATCTCATGAAGAAACCGAGCAT	525
Qy	642	AGTTACACAGCTGGAGAGACAGTTTCCAAAGACTCTAATCAGGATATCGAGAGAGATGATG	701
Db	526	AGTTACACAGCTGGAGAGACAGTTTCCAAAGACTCTAATCAGGATATCGAGAGAGATGATG	585
Qy	702	TCTGAGCAGGAAAATCCAGATTCAGTGAACCAAGTAGTGAAGATGAAGATTTGACCAAT	761
Db	586	TCTGAGCAGGAAAATCCAGATTCAGTGAACCAAGTAGTGAAGATGAAGATTTGACCAAT	645
Qy	762	GATCAGATGATGTAAACATACCAAGTCTTATGAGGAAACCAAGCAGTATATGAACCTCTAGAA	821
Db	646	GATCAGATGATGTAAACATACCAAGTCTTATGAGGAAACCAAGCAGTATATGAACCTCTAGAA	705
Qy	822	AATGAAGGATAGAAATCACAAGTAACCTG	852
Db	706	AATGAAGGATAGAAATCACAAGTAACCTG	736

```

RESULT 11
US-09-949-016-15189/c
; Sequence 15189, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15189
; LENGTH: 74730
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(74730)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15189

```

```

Qy 561 GATGAGTTCCTTATGGCGACTGATGTAGATGATTTTGAGACCCCTGGAACCTGGAAGTA 620
Db 29846 GATGAGTTCCTTATGGCGACTGATGTAGATGATTTTGAGACCCCTGGAACCTGGAAGTA 29787

Qy 621 TCTCATGAAG 630
Db 29786 TCTCATGAAG 29777

RESULT 12
US-09-949-016-3447/c
; Sequence 3447, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001301
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3447
; LENGTH: 956
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3447

Query Match 4.9%; Score 115; DB 4; Length 956;
Best Local Similarity 100.0%; Pred. No. 6.4e-45;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 516 TTGCAACAAGAAGATGGACCCACAGGAGAACCAACAAGAGATGATGAGTTCCTTATG 575
Db 462 TTGCAACAAGAAGATGGACCCACAGGAGAACCAACAAGAGATGATGAGTTCCTTATG 403

Qy 576 GCGACTGATGTAGATGATAGTTTGAGACCCCTGGAACCTGGAATCTCTCATGAAG 630
Db 402 GCGACTGATGTAGATGATAGTTTGAGACCCCTGGAACCTGGAATCTCTCATGAAG 348

```

```

RESULT 13
US-09-949-016-16453/c
? Sequence 16453, Application US/09949016
? Patent No. 6812339
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: C0001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 16453
? LENGTH: 5746
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-16453
Query Match 3.8%; Score 88; DB 4; Length 5746;

```

Best Local Similarity 100.0%; Pred. No. 6e-32; DB 4; Length 601;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1073 GGGAAAAATTGAGAACGAGTGAATGCATTTAAAGAACTAGTAGCGAAATACCTCTCAGAG 1132  
Db 696 GGGAAAAATTGAGAACGAGTGAATGCATTTAAAGAACTAGTAGCGAAATACCTCTCAGAG 637

Qy 1133 TCCACGAGCAAGATATGGGAGGCGCAG 1160  
Db 636 TCCACGAGCAAGATATGGGAGGCGCAG 609

## RESULT 14

US-09-949-016-122372/c  
; Sequence 122372, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122372  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-122372

Query Match 3.3%; Score 77; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.2e-26;  
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 425 AGACCCAGAAATATCGAAGTGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAAT 484  
Db 268 AGAACCCAGAAATATCGAAGTGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAAT 209

Qy 485 GGTACACGCAAGCAATG 501  
Db 208 GGTACACGCAAGCAATG 192

## RESULT 15

US-09-949-016-122387/c  
; Sequence 122387, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122387  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-122387

## US-09-949-016-122387

Query Match 3.1%; Score 73; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 9.8e-25;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 262 TAGAAAACTAGGAATCTATGATGCTGATGGTGGATGAGATTTTGTGATGGATGATGCCA 321  
Db 86 TAGAAAACTAGGAATCTATGATGCTGATGGTGGATGAGATTTTGTGATGGATGATGCCA 27

Qy 322 AAGTTTATTAGG 334  
Db 26 AAGTTTATTAGG 14

## RESULT 16

US-09-949-016-236/c  
; Sequence 236, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 236  
; LENGTH: 2208  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-236

Query Match 2.5%; Score 59; DB 4; Length 2208;  
Best Local Similarity 100.0%; Pred. No. 5.2e-18;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2239 GGCATCCGGAACCTGACACCAACAGAGAGCGAGCGCTTCCAGCAATTTAGCATGAATC 2297  
Db 59 GGCATCCGGAACCTGACACCAACAGAGAGCGAGCGCTTCCAGCAATTTAGCATGAATC 1

## RESULT 17

US-09-949-016-122345/c  
; Sequence 122345, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122345  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-122345

Query Match 1.9%; Score 45; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.8e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 943 CAGAAACAATAGAAAAACAGATGATCCAGAAACAAAAAGCAAAAG 987  
|||  
DB 212 CAGAAACAATAGAAAAACAGATGATCCAGAAACAAAAAGCAAAAG 168  
|||

RESULT 18  
US-09-949-016-13283  
; Sequence 13283, Application US/09949016  
; Patent No. 6812319  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13283  
; TYPE: DNA  
; LENGTH: 9984  
; ORGANISM: Human  
US-09-949-016-13283

Query Match 0.9%; Score 22; DB 4; Length 9984;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 CAGCAGCAGCAGCGGCTCCGGC 68  
|||  
DB 920 CAGCAGCAGCAGCGGCTCCGGC 941  
|||

RESULT 19  
US-09-614-034-188  
; Sequence 188, Application US/09614034  
; Patent No. 6489307  
; GENERAL INFORMATION:  
; APPLICANT: PHILLIPS, M. IAN  
; APPLICANT: ZHANG, YUAN  
; TITLE OF INVENTION: ANTISENSE COMPOSITIONS TARGETED TO BETAL-ADRENOCEPTOR-SPECIFIC MR  
; TITLE OF INVENTION: METHODS OF USE  
; FILE REFERENCE: 4300.013900  
; CURRENT APPLICATION NUMBER: US/09/614,034  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 09/152,717  
; PRIOR FILING DATE: 1998-09-14  
; PRIOR APPLICATION NUMBER: PCT/US99/21007  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 188  
; LENGTH: 1845  
; TYPE: DNA  
; ORGANISM: UNKNOWN  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC OLIGONUCLEOTIDE  
; NAME/KEY: misc feature  
; LOCATION: (1828)..(1828)  
; OTHER INFORMATION: WHERE N = A, T, C OR G  
US-09-614-034-188

Query Match 0.9%; Score 21; DB 4; Length 1845;  
Best Local Similarity 100.0%; Pred. No. 9.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CAGCAGCGGCTCCGGCAGCGG 73  
|||  
DB 93 CAGCAGCGGCTCCGGCAGCGG 113  
|||

RESULT 20  
US-09-859-604-10/c  
; Sequence 10, Application US/09859604  
; Patent No. 6835370  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R., Suzanne M  
; APPLICANT: de la Monte, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Location (-1)  
; OTHER INFORMATION: oligonucleotide  
US-09-859-604-10

Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AATGCCCCAGCGTAAGAATG 30  
|||  
DB 20 AATGCCCCAGCGTAAGAATG 1  
|||

RESULT 21  
US-09-859-604-11/c  
; Sequence 11, Application US/09859604  
; Patent No. 6835370  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M  
; APPLICANT: de la Monte, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Location (-6)  
; OTHER INFORMATION: oligonucleotide  
US-09-859-604-11

Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 0.9%; Score 20; DB 4; Length 342;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 26  
US-09-949-016-3254  
; Sequence 3254, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3254  
; LENGTH: 1106  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-3254

Query Match 0.9%; Score 20; DB 4; Length 1106;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
|||||  
Db 343 GCAACAGCAGCAGCAGCGGC 362

## RESULT 27

US-09-949-016-1812  
; Sequence 1912, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1812  
; LENGTH: 1116  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-1812

Query Match 0.9%; Score 20; DB 4; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
|||||  
Db 353 GCAACAGCAGCAGCAGCGGC 372

## RESULT 28

US-09-949-016-14996  
; Sequence 14996, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14996  
; LENGTH: 8841  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14996

; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14996  
; LENGTH: 8841  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14996

Query Match 0.9%; Score 20; DB 4; Length 8841;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
|||||  
Db 6077 GCAACAGCAGCAGCAGCGGC 6096

## RESULT 29

US-09-949-016-11878  
; Sequence 11878, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11878  
; LENGTH: 11668  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11878

Query Match 0.9%; Score 20; DB 4; Length 11668;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
|||||  
Db 8971 GCAACAGCAGCAGCAGCGGC 8990

## RESULT 30

US-09-949-016-13554  
; Sequence 13554, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13554  
; LENGTH: 11668  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13554

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13554  
; LENGTH: 11713  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13554

Query Match 0.9%; Score 20; DB 4; Length 11713;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
|||||  
Db 8949 GCAACAGCAGCAGCAGCGGC 8968

## RESULT 31

US-09-949-016-17029  
; Sequence 17029, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17029  
; LENGTH: 47284  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17029

Query Match 0.9%; Score 20; DB 4; Length 47284;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 TGATGCCAAAGTTTATTAG 333  
|||||  
Db 20734 TGATGCCAAAGTTTATTAG 20753

## RESULT 32

US-09-949-016-12762/c  
; Sequence 12762, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12762

; LENGTH: 117937  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12762

Query Match 0.9%; Score 20; DB 4; Length 117937;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
|||||  
Db 50840 GAACAAAAGCAAAAGTTAA 50821

## RESULT 33

US-09-949-016-15775/c  
; Sequence 15775, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15775  
; LENGTH: 117937  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15775

Query Match 0.9%; Score 20; DB 4; Length 117937;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
|||||  
Db 50840 GAACAAAAGCAAAAGTTAA 50821

## RESULT 34

US-09-513-999C-35281  
; Sequence 35281, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59, US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 35281  
; LENGTH: 148  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-513-999C-35281

Query Match 0.8%; Score 19; DB 4; Length 148;  
Best Local Similarity 100.0%; Pred. No. 88;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 AACAGCAGCAGCAGCGCT 63  
Db 20 AACAGCAGCAGCAGCGCT 38

## RESULT 35

US-09-248-796A-13462/c  
; Sequence 13462, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 13462  
; LENGTH: 309  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-13462

Query Match 0.8%; Score 19; DB 4; Length 309;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1004 TAACTTTTAAATAAATTT 1022  
Db 168 TAACTTTTAAATAAATTT 150

## RESULT 36

US-09-949-016-61822/c  
; Sequence 61822, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 61822  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-61822

Query Match 0.8%; Score 19; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 CAGCAGCAGCAGCGCTCC 65  
Db 109 CAGCAGCAGCAGCGCTCC 91

## RESULT 37

US-09-248-796A-4406/c  
; Sequence 4406, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 4406  
; LENGTH: 951  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-4406

Query Match 0.8%; Score 19; DB 4; Length 951;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1004 TAACTTTTAAATAAATTT 1022  
Db 810 TAACTTTTAAATAAATTT 792

## RESULT 38

US-09-949-016-1792/c  
; Sequence 1792, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1792  
; LENGTH: 1599  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-1792

Query Match 0.8%; Score 19; DB 4; Length 1599;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 CAGCAGCAGCAGCGCTCC 65  
Db 963 CAGCAGCAGCAGCGCTCC 945

## RESULT 39

US-08-339-214-25/c  
; Sequence 25, Application US/08339214  
; Patent No. 6348334  
; GENERAL INFORMATION:  
; APPLICANT: Nagata, Shigikazu  
; APPLICANT: Suda, Takashi  
; APPLICANT: Takahashi, Tomoniro  
; APPLICANT: Nakamura, No. 634833410



;; TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA  
;; TITLE OF INVENTION: Encoding the Same  
;; NUMBER OF SEQUENCES: 103  
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
;; STREET: P.O. Box 747  
;; CITY: Falls Church  
;; STATE: Virginia  
;; COUNTRY: USA  
;; ZIP: 22040-0747

## COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/339,214  
;; FILING DATE: 10-NOV-1994  
;; CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

;; NAME: Murphy Jr. Gerald M.  
;; REGISTRATION NUMBER: 28,977  
;; REFERENCE/DOCKET NUMBER: 1110-139P  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-205-8000  
;; TELEFAX: 703-205-8050

## INFORMATION FOR SEQ ID NO: 25:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1623 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: cDNA to mRNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: rat  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 74...907  
;; OTHER INFORMATION: /label= Figs\_2-3  
US-08-339-214-25

Query Match 0.8%; Score 19; DB 3; Length 1623;  
Best Local Similarity 100.08; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 812 ACCTCTAGAAATGAAGG 830  
|||||  
Db 1052 ACCTCTAGAAATGAAGG 1034

## RESULT 40

US-09-149-476-110/c  
;; Sequence 110, Application US/09149476  
;; Patent No. 6420526  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: 186 Human Secreted proteins  
;; FILE REFERENCE: PZ002P1  
;; CURRENT APPLICATION NUMBER: US/09/149,476  
;; CURRENT FILING DATE: 1998-09-08  
;; EARLIER APPLICATION NUMBER: PCT/US98/04493  
;; EARLIER FILING DATE: 1998-03-06  
;; EARLIER APPLICATION NUMBER: 60/040,162  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,333  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/038,621  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,626  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,334

;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,336  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,163  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/047,600  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,615  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,597  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,502  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,633  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,583  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,617  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,618  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,503  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,592  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,581  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,584  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,500  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,587  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,492  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,598  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,613  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,582  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,596  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,612  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,632  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,601  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/043,580  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,568  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,314  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,569  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,311  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,671  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,674  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,669  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,312  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,313  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,672  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,315  
;; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,845  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,892  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,761  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/047,595  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,599  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,588  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,585  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,586  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,590  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,594  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,589  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,593  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,614  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,578  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,576

; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/047,501  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,670  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/056,632  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,664  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,876  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,881  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,909  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,875  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,862  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,887  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,908  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/057,650  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/056,884  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,669  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/049,610  
; EARLIER FILING DATE: 1997-06-13  
; EARLIER APPLICATION NUMBER: 60/061,060  
; EARLIER FILING DATE: 1997-10-02  
Query Match 0.8%; Score 19; DB 3; Length 1751;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1836 AAAGTCTCTTCTGCGCTG 1854  
Db 692 AAAGTCTCTTCTGCGCTG 674  
RESULT 41  
US-09-949-016-5583  
; Sequence 5583, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ IDS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5583  
; LENGTH: 1821  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-5583

Query Match 0.8%; Score 19; DB 4; Length 1821;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 36 AGCAGCGGCACAGCAGCA 54  
 |||||  
 Db 437 AGCAGCGGCACAGCAGCA 455

## RESULT 42

US-09-620-312D-49/c  
 ; Sequence 49, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunqing

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: John Tillinghast

; APPLICANT: Drmanac, Radjoe T.

; TITLE OF INVENTION: No. 659662el Nucleic Acids and

; FILE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: pt\_FL\_genes Version 1.0

; SEQ ID NO 49

; LENGTH: 2286

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (69)..(1727)

US-09-620-312D-49

Query Match 0.8%; Score 19; DB 4; Length 2286;

Best Local Similarity 100.0%; Pred. No. 88;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1836 AAAGTCTCTCTCGCTG 1854

|||

Db 1839 AAAGTCTCTCTCGCTG 1821

## RESULT 43

US-09-149-476-264/c

; Sequence 264, Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

; FILE REFERENCE: P2002F1

; CURRENT APPLICATION NUMBER: US/09/149,476

; CURRENT FILING DATE: 1998-09-08

; EARLIER APPLICATION NUMBER: PCT/US98/04493

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,333

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/038,621

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,626

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,334

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,336

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,163

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/047,600

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,615

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,502

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,633

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,583

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,617

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,618

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,503

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,592

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,581

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,584

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,500

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,587

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,492

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,598

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,613

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,582

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,596

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,612

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,632

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/043,601

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/043,580

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,568

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,314

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,569

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,311

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,671

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,674

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,669

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,312

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,313

```

; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23

```

```

; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 0.8% Score 19; DB 3; Length 2377;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1836 AAAGTCTCTCTCTGCGCTG 1854
Db 1676 AAAGTCTCTCTCTGCGCTG 1658

RESULT 44
US-09-949-016-2642
; Sequence 2642, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2642
; LENGTH: 3462
; TYPE: DNA
; ORGANISM: Human

```

## US-09-949-016-2642

Query Match 0.8%; Score 19; DB 4; Length 3462;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1777 CTTTAGAAGAACTGGAA 1795  
| | | | | | | | | | | | | | | | | | | | | |  
Db 2966 CTTTAGAAGAACTGGAA 2984

## RESULT 45

US-09-949-016-633  
; Sequence 633, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 633  
; LENGTH: 3482  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(3482)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-633

Query Match 0.8%; Score 19; DB 4; Length 3482;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1777 CTTTAGAAGAACTGGAA 1795  
| | | | | | | | | | | | | | | | | | | | | |  
Db 2975 CTTTAGAAGAACTGGAA 2993

## RESULT 46

US-09-949-016-13534/c  
; Sequence 13534, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13534  
; LENGTH: 6625  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13534

Query Match 0.8%; Score 19; DB 4; Length 6625;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 CAGCAGCAGCGGCTCC 65  
| | | | | | | | | | | | | | | | | | | | | |  
Db 3989 CAGCAGCAGCGGCTCC 3971

## RESULT 47

US-09-949-016-12375  
; Sequence 12375, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12375  
; LENGTH: 7460  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12375

Query Match 0.8%; Score 19; DB 4; Length 7460;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1777 CTTTAGAAGAACTGGAA 1795  
| | | | | | | | | | | | | | | | | | | | | |  
Db 4966 CTTTAGAAGAACTGGAA 4984

## RESULT 48

US-09-949-016-14384  
; Sequence 14384, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14384  
; LENGTH: 7462  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14384

Query Match 0.8%; Score 19; DB 4; Length 7462;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1777 CTTTGAAGAAGAACTGGAA 1795  
Db 4966 CTTTGAAGAAGAACTGGAA 4984

RESULT 49  
US-09-949-016-15619  
; Sequence 15619, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15619  
; LENGTH: 7700  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15619

Query Match 0.8%; Score 19; DB 4; Length 7700;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1777 CTTTGAAGAAGAACTGGAA 1795  
Db 6898 CTTTGAAGAAGAACTGGAA 6916

RESULT 50  
US-09-949-016-17325  
; Sequence 17325, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17325  
; LENGTH: 28843  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(28843)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17325

Query Match 0.8%; Score 19; DB 4; Length 28843;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 AGCAGCGGCACAGCAGCA 54

Db 18759 AGCAGCGGCACAGCAGCA 18777

RESULT 51  
US-09-949-016-11868/c  
; Sequence 11868, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11868  
; LENGTH: 300598  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(300598)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-11868

Query Match 0.8%; Score 19; DB 4; Length 300598;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1072 GCGGAAAAATTGAGGAAGC 1090  
Db 249442 GCGGAAAAATTGAGGAAGC 249424

RESULT 52  
US-09-949-016-14588/c  
; Sequence 14588, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14588  
; LENGTH: 302604  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(302604)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14588

Query Match 0.8%; Score 19; DB 4; Length 302604;  
Best Local Similarity 100.0%; Pred. No. 86;

```
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1072 GGGGAAAAATTGAGGAAGC 1090
    |||||
Db 299448 GGGGAAAAATTGAGGAAGC 299430

RESULT 53
US-09-949-016-14589/c
; Sequence 14589, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14589
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14589

Query Match 0.8%; Score 19; DB 4; Length 302604;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1072 GGGGAAAAATTGAGGAAGC 1090
    |||||
Db 299448 GGGGAAAAATTGAGGAAGC 299430

RESULT 54
US-09-949-016-17119/c
; Sequence 17119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17119
; LENGTH: 308362
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(308362)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17119
```

```
Query Match 0.8%; Score 19; DB 4; Length 308362;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1072 GGGGAAAAATTGAGGAAGC 1090
    |||||
Db 299264 GGGGAAAAATTGAGGAAGC 299246

RESULT 55
US-09-949-016-12777/c
; Sequence 12777, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12777
; LENGTH: 828152
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(828152)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12777

Query Match 0.8%; Score 19; DB 4; Length 828152;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGCTGATGG 298
    |||||
Db 816467 ATGATGCTGATGCTGATGG 816449

RESULT 56
US-09-859-604-13
; Sequence 13, Application US/09859604
; Patent No. 6835370
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M
; APPLICANT: Deutch, Alan H
; APPLICANT: Ghanbari, Hossein A
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 CIP
; CURRENT APPLICATION NUMBER: US/09/859,604
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Sense
; OTHER INFORMATION: Oligonucleotide
US-09-859-604-13
```

```
Query Match      0.8%; Score 18; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TGCAATGCCCGCGTAA 25
Db 5 TGCAATGCCCGCGTAA 22

RESULT 57
US-09-313-294A-3024
; Sequence 3024, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 3024
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700282183H1
; NAME/KEY: unsure
; LOCATION: 87, 108, 146, 200, 239
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3024

Query Match      0.8%; Score 18; DB 4; Length 286;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 CAACAGCAGCAGCGCG 61
Db 181 CAACAGCAGCAGCGCG 198

RESULT 58
US-08-482-842B-2
; Sequence 2, Application US/08482842B
; Patent No. 5910480
; GENERAL INFORMATION:
; APPLICANT: KOWAN, Ahmet
; APPLICANT: CHASSIN, Dorine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN,
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF THIS PROTEIN AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION CONTAINING SUCH, DNA CODING FOR
; TITLE OF INVENTION: SAID PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,842B
; FILING DATE: 07-JUN-1995
```

```
CLASSIFICATION: 514
PRIOR APPLICATION DATA: FR 94-07191
APPLICATION NUMBER: 13-JUN-1994
FILING DATE: 13-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 017753-058
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 319 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-482-842B-2

Query Match      0.8%; Score 18; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 984 AAAGTTAAGAAAGAG 1001
Db 88 AAAGTTAAGAAAGAG 105

RESULT 59
US-09-513-999C-10601
; Sequence 10601, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; PATENT NO. 6783961
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 10601
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 440
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-10601

Query Match      0.8%; Score 18; DB 4; Length 452;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1495 GCTTCATCCTGAGGCAC 1512
Db 200 GCTTCATCCTGAGGCAC 217

RESULT 60
US-09-270-767-10408
; Sequence 10408, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
```



; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10408  
; LENGTH: 529  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-10408

Query Match 0.8%; Score 18; DB 4; Length 529;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCG 60  
Db 136 GCAACAGCAGCAGCAGCG 153

RESULT 61  
US-09-918-686-4/c  
; Sequence 4, Application US/09918686  
; Patent No. 6475739  
; GENERAL INFORMATION:  
; APPLICANT: Brunkow, Mary  
; APPLICANT: Proll, Sean  
; APPLICANT: Paepker, Bryan  
; APPLICANT: Staehling-Hampton, Karen  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: GENOMIC DELETIONS  
; FILE REFERENCE: 240083.515  
; CURRENT APPLICATION NUMBER: US/09/918,686  
; CURRENT FILING DATE: 2001-07-30  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 557  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-918-686-4

Query Match 0.8%; Score 18; DB 4; Length 557;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1830 AAGCCAAAGTCTCTTC 1847  
Db 265 AAGCCAAAGTCTCTTC 248

RESULT 62  
US-09-621-976-16896/c  
; Sequence 16896, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 16896  
; LENGTH: 561  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-16896

Query Match 0.8%; Score 18; DB 4; Length 561;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2100 AAGCAAGGCTGCAAGATT 2117  
Db 349 AAGCAAGGCTGCAAGATT 332

## RESULT 63

US-09-949-016-36982/c  
; Sequence 36982, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36982  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-36982

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1909 AAGCAAGCAAGATGAAA 1926  
Db 490 AAGCAAGCAAGATGAAA 473

## RESULT 64

US-09-949-016-40425/c  
; Sequence 40425, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40425  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-40425

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1602 CTGGGGATGCCATGCAG 1619  
Db 414 CTGGGGATGCCATGCAG 397

```
RESULT 65
US-09-949-016-42135
; Sequence 42135, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42135
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-42135

Query Match          0.8%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297
Db 320 ATGATGCTGATGGTGATG 337

RESULT 66
US-09-949-016-42136
; Sequence 42136, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42136
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-42136

Query Match          0.8%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297
Db 280 ATGATGCTGATGGTGATG 297

RESULT 67
US-09-949-016-42137
; Sequence 42137, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42137
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-42137

Query Match          0.8%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297
Db 208 ATGATGCTGATGGTGATG 225

RESULT 68
US-09-949-016-42138
; Sequence 42138, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42138
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-42138

Query Match          0.8%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297
Db 222 ATGATGCTGATGGTGATG 239

RESULT 69
US-09-949-016-42139
; Sequence 42139, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42139  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42139

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297  
Db 205 ATGATGCTGATGGTGATG 222

## RESULT 70

US-09-949-016-42140  
; Sequence 42140, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42140  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42140

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297  
Db 186 ATGATGCTGATGGTGATG 203

## RESULT 71

US-09-949-016-53139/c  
; Sequence 53139, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53139  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-53139

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1772 AAAGTCTTTAGAAAGAAA 1789  
Db 214 AAAGTCTTTAGAAAGAAA 197

## RESULT 72

US-09-949-016-67303  
; Sequence 67303, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 67303  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-67303

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred.No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 AGCAGCAGCAGCGGCTCC 65  
Db 41 AGCAGCAGCAGCGGCTCC 58

## RESULT 73

US-09-949-016-69417/c  
; Sequence 69417, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 69417  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-69417

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 GAACCTGAAGTATCTCAT 626  
|||||  
Db 280 GAACCTGAAGTATCTCAT 263

## RESULT 74

US-09-949-016-69418/c  
; Sequence 69418, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 69418  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-69418

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 GAACCTGAAGTATCTCAT 626  
|||||  
Db 372 GAACCTGAAGTATCTCAT 355

## RESULT 75

US-09-949-016-69419/c  
; Sequence 69419, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 69419  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-69419

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 GAACCTGAAGTATCTCAT 626  
|||||  
Db 463 GAACCTGAAGTATCTCAT 446

Search completed: March 25, 2005, 22:35:18  
Job time : 417 secs

This Page Blank (uspto)